Université de Montréal

Trois essais sur le comportement optimal dans les marchés financiers

par

Marcel Rindisbacher

Département des sciences économiques

Faculté des arts et des sciences

Thèse présentée à la Faculté des études supérieures

en vue de l'obtention du grade de

PhilosophiæDoctor (Ph.D.)

en sciences économiques

Avril 2000

Université de Montréal
Faculté des études supérieures

Cette thèse intitulée:

Trois essais sur le comportement optimal dans les marchés financiers

présentée par :

Marcel Rindisbacher

a été évaluée par un jury composé des personnes suivantes:

Président-rapporteur :	Jean-Marie Dufour (Gérard Gaudet)
Directeur de recherche :	René Garcia
Membre du jury :	Nour Medahhi (Benoit Perron)
Examinateur externe :	Nizar Touzi
	(Université Paris I Panthéon-Sorbonne)
	Phelim Boyle
	(University of Waterloo)
Représentant du doyen :	

Thèse acceptée le : 29 avril 2000

Sommaire

Étant donné que la création récente d'un grand nombre de fonds mutuels et leur importance comme moyen d'épargne dans l'économie, il semble surprenant que la gestion actuelle des fonds d'actifs financiers en pratique est presque exclusivement basée sur des modèles financiers statiques comme le modèle d'évaluation des actifs financiers ou des modèles à facteurs statiques. Ceci implique que tous les conseils d'investissement basés sur de tels modèles ne dépendent pas de l'horizon d'investissment, ce qui ne correspond pas à la composition des portefeuilles observée en réalité où des effets d'âge sont clairement présents.

Il semble alors important de développer une théorie d'allocation de portefeuille dynamique qui peut servir comme base pour la mise en oeuvre des systèmes de support pour les décisions des gestionaires de fonds semblables à ceux utilisés pour la valorisation et la couverture des produits dérivées. Théoriquement, le problème de choix de portefeuille dynamique a été étudié dans les années soixante dix par Samuelson et Merton. Le résultat principal de leurs travaux, comparé à la demande de portefeuille dans un modèle statique, est que le comportement optimal de l'agent est tel que, à part une composante basée sur des primes de risque instantanées, l'investisseur essaie de se couvrir contre les fluctuations futures dans l'ensemble des opportunités d'investissement. La caractérisation des stratégies optimales dans ces modèles est faite par des éléments de la théorie du contrôle optimal, mais peu de travaux ont été éffectués pour la mise en pratique

de ces résultats. L'une des raisons est que la mise en oeuvre de cette méthode avec beaucoup de variables d'état (qui est typique pour le choix de portefeuille) est difficile.

Le présent travail tente de faire des contributions théoriques et pratiques pour la compréhension approfondie de la gestion de portefeuille dans un contexte dynamique. Nous dérivons une décomposition probabiliste de la politique de portefeuille qui admet une identification des éléments déterminants de la composition du portefeuille optimale de l'investisseur.

D'abord nous présentons des résultats sur les stratégies de portefeuille qui aident à identifier la dépendance de la demande pour les actifs financiers du flux d'information dans l'économie. En particulier, nous considérons le comportement optimal d'un investisseur qui a acqui des connaissances anticipatives relative au flux d'information publique. Nous analysons en détail les types d'information de l'initié qui permettent de réaliser des possibilités d'arbitrage et dérivons la stratégie de portefeuille optimale d'un tel investisseur.

Ensuite nous dérivons une méthode pour la solution numérique du problème de choix de portefeuille basée sur des simulations de Monte Carlo et de solution numérique d'équations différentielles stochastiques, comme dans les méthodes utilisées pour la valorisation des options. Notre méthode est générale et est avantageuse par rapport à la programmation dynamique numérique si le flux d'information dans l'économie est généré par un grand nombre de variables d'état.

Finalement, nous faisons une analyse d'erreur approfondie de notre méthode numérique, et dérivons des procedures efficaces au niveau de la vitesse de convergence et de la correction du biais asymptotique de deuxième ordre. Notre approche admet l'analyse d'erreur pour une classe de problèmes différents, comme la valorisation et la couverture des options, la détermination de la volatilité des actifs financiers risqués dans un modèle d'équilibre général ou l'analyse d'erreur des processus en temps discret comme approximations de diffusions.

Résumé

La thèse intitulée “Trois essais sur le comportement optimal dans les marchés financiers” contient trois articles sur les stratégies de portefeuille optimales des agents dans des marchés financiers.

Depuis les années soixante dix, les marchés financiers ont connu une révolution relative aux titres qui y sont échangés. Cette révolution a été causé par la découverte de la formule de Black and Scholes pour l'évaluation des options. La compréhension approfondie de la valorisation d'actifs financiers a initié la création de nouveaux actifs financiers et de nouvelles bourses où de tels titres sont échangés. Au niveau théorique, l'importance de ce résultat vient de l'idée de réplication d'un portefeuille qui, ajoutée à l'idée d'absence d'arbitrage, est la base de toute valorisation dans la théorie financière moderne. Les techniques mathématiques utilisées pour démontrer ce résultat sont basées sur des solutions d'équations aux dérivées partielles (EDP) et sont alors liées à des résultats probabilistes tels que découverts dans la célèbre formule de Feynman et Kac et du calcul stochastique. Cette innovation technique a aidé à démontrer certain résultats dans une plus grande généralité et a donné naissance à une nouvelle formation d'ingénieur financier.

Vu ces innovations majeurs dans l'évaluation des produits derivés, il est surprenant que, encore aujourd'hui, ces changements dans la théorie financière récente ont eu relativement peu d'influence sur la gestion de portefeuille optimale dans un contexte dynamique. Ceci est plus étonant si l'on considère les changements

majeurs des consommateurs dans leur comportement d'épargne qui est révélé par la création de nombreux fonds mutuels. Le valeur et le volume de ces fonds montrent l'importance d'une compréhension approfondie de la gestion de portefeuille optimale. Encore aujourd'hui, les gestionnaires de fonds utilisent des méthodes moyenne-variance introduites par Markovitz dans les années cinquante. Mais ces principes de gestion de portefeuille qui portent surtout sur l'idée de diversification sont fondés sur un modèle financier statique avec des préférences de type moyenne-variance et ne peuvent donc pas apporter des réponses à des questions, telles que de l'effet de l'horizon d'investissement sur les stratégies optimales. D'autre part, les gestionnaires de fond donnent des conseils d'investissement qui dépendent du temps sans s'appuyer sur une base théorique. Ceci est d'autant plus surprenant que l'idée de réplication et couverture qui est utilisée dans l'évaluation des produits derivés nécessite des considérations dynamiques et est dans ce sense très proche de l'idée de gestion optimale de portefeuille. Une stratégie de portefeuille optimale réplique la consommation et/ou la richesse terminale optimale de l'investisseur. Si on connaît alors la consommation optimale de l'agent, le problème de trouver une stratégie de portefeuille optimale consiste à trouver une composition d'actifs financiers qui produise des flux tels que dans tous les scénarios possibles et dans tous les moments avant l'horizon d'investissement, l'agent a les moyens nécessaires de financer sa consommation optimale.

Dans un cadre dynamique, la stratégie de portefeuille optimale était décrite pour la première fois par Merton en soixante douze. Il a réalisé qu'une stratégie de portefeuille optimal contient une partie qui correspond à la demande moyenne-variance mais aussi une partie de couverture contre les fluctuations futures dans l'ensemble des possibilités d'investissement.

Dans les trois articles suivants, nous présentons plusieurs extensions du problème de portefeuille. Dans un premier article nous analysons l'effet d'information anticipative sur la stratégie de portefeuille. Ce papier contient des résultats importants pour la construction de modèles d'équilibre avec des investisseurs informés. La généralité des résultats permet aussi d'analyser sous quelles conditions les marchés financiers sont des mécanismes efficaces pour l'agrégation d'information dans l'économie et aide à comprendre le comportement d'un initié dans un marché compétitif.

Dans un deuxième article nous présentons une nouvelle méthode de calcul des stratégies de portefeuille. Cette méthode est basée sur des méthodes purement probabilistes et est alors avantageuse relativement à la programmation dynamique standard si le nombre de variables d'état est grand. Nous illustrons notre méthode avec plusieurs exemples. Notre représentation de la stratégie de portefeuille optimale permet aussi de mieux comprendre la structure de la demande de portefeuille et aide à identifier des cas particuliers pour lesquels il existe des solutions explicites à la demande d'actifs risqués.

Le dernier article fournit des expressions pour les lois asymptotiques présentées dans le deuxième article. Ces expressions sont importantes pour analysér la précision de différentes méthodes d'approximation. Ces expressions permettent de construire des tests asymptotiques et des intervalles de confiance pour les différentes parties du portefeuille optimal. Etant donné un budget de temps de calcul préscrit, les résultats permettent de planifier une mise en oeuvre efficace de la procédure. Nous montrons aussi des moyens d'augmenter la vitesse de convergence et dérivons à partir des loi asymptotiques des procédures de correction de biais du deuxième ordre.

Premier article

Dans le premier article intitulée "Insider Information, Arbitrage and Optimal Portfolio and Consumption Policies", nous ajoutons des initiés dans le modèle financier introduit par Samuelson (1969) et Merton (1971). Nous montrons que si l'horizon d'investissement de ces initiés est postérieur à la dissipation de leur avantage informationnel, ils ont des possibilités d'arbitrage si et seulement si leur information est si précise qu'elle contient des événements qui sont de probabilité nulle sachant l'information publique. Pour un horizon d'investissement plus court ou en l'absence de tels évènements, nous dérivons des expressions explicites pour les politiques de consommation et d'investissement à tous les horizons d'investissement, qui permettent d'analyser comment les politiques de consommation et d'investissement sont affectées par l'information de l'initié. Nous montrons que les stratégies optimales ne révèlent jamais toute l'information et que l'information d'initié est sans valeur si et seulement si elle est indépendante de l'information publique. Nous montrons que les possibilités d'arbitrage sont telles que les initiés peuvent couvrir toutes les politiques de consommation. Par conséquent, le problème d'investissement de Merton avec des préférences convexes du type von Neumann-Morgenstern n'a pas de solution pour des horizons d'investissement qui se terminent après la résolution de l'incertitude concernant le signal. Ce problème peut être évité si on ajoute au signal de l'initié un bruit. Comme dans ce cas les non-initiés n'apprennent jamais le vrai signal, une telle solution néglige des aspects importants de l'information de l'initié. Pour un initié, qui a une information additionnelle générée par un signal avec loi discrète, nous montrons que la valorisation des options mesurables par rapport à l'information publique n'est pas affectée s'il n'y a pas de possibilités d'arbitrage. Au contraire, pour un initié avec une information additionelle générée par un signal avec loi continue, toutes les options sont sans valeur.

Deuxième article

Le deuxieème article intitulée "A Monte Carlo Method for Optimal Portfolios" établit des résultats nouveaux sur (i) la structure des portefeuilles optimaux, (ii) le comportement des termes de couverture et (iii) les méthodes numériques de simulation en la matière. Le fondement de notre approche repose sur l'obtention de formules explicites pour les dérivées de Malliavin de processus de diffusion, formules qui simplifient leur simulation numérique et facilitent le calcul des composantes de couverture des portefeuilles optimaux. Une de nos procédures utilise une transformation des processus sous-jacents qui élimine les intégrales stochastiques de la représentation des dérivées de Malliavin et assure l'existence d'une approximation faible exacte. Cette transformation améliore alors la performance des méthodes de Monte-Carlo lors de l'implémentation numérique des politiques de portefeuille dérivées par des méthodes probabilistes. Notre approche est flexible et peut être utilisée même lorsque la dimension de l'espace des variables d'état sous-jacentes est grande. Cette méthode est appliquée dans le cadre de modèles bivariés et trivariés dans lesquels l'incertitude est décrite par des mouvements de diffusion pour le prix de marché du risque, le taux d'intérêt et les autres facteurs d'importance. Après avoir calibré le modèle aux données nous examinons le comportement du portefeuille optimal et des composantes de couverture par rapport aux paramètres tels que l'aversion au risque, l'horizon d'investissement, le taux d'intérêt et le prix de risque du marché. Nous démontrons l'importance des demandes de couverture. L'aversion au risque et l'horizon d'investissement émergent comme des facteurs déterminants qui ont un impact substantiel sur la taille du portefeuille optimal et sur ses propriétés économiques.

Finalement nous analysons dans un modèle avec deux actifs risqués et un actif sans risque le comportement des différents composantes du portefeuille par rapport à la corrélation des prix du risque et à la corrélation des rendements des actifs risqués. Nous montrons qu'un actif risqué est choisi comme couverture contre les futures fluctuations dans le taux d'intérêt. Puis, comme cet actif a une corrélation positive avec le taux, une telle couverture est atteinte par une position en compte dans cet actif. Étant donné cette position en compte, l'autre actif est alors utilisé comme couverture contre le risque apporté par cette position de couverture. Il est détenu dans une position à découverte si les rendements sont négativement corrélés ou en compte si la corrélation des rendements est négative. Les positions opposées augmentent avec le valeur absolue de la corrélation entre les rendements. Un comportement similaire est analysé pour la partie couverture contre le risque représenté par les futures fluctuations du prix du risque.

Troisième article

Dans le dernier article intitulée "Asymptotic Properties of Optimal Portfolio Estimators" nous montrons d'abord comment notre méthode de résolution du problème de portfeuille peut être, en élargissant l'espace des variables d'état, définit comme un probléme d'estimation de l'espérance d'une fonction du valeur terminale d'une équation différentielle stochastique comme dans l'évaluation des options exotiques où le payoff est dépendant de la trajectoire des actifs sous-jacents. Ensuite, nous dérivons d'abord les lois asymptotiques pour les variables d'états avec ou sans transformation qui élimine les intégrales stochastiques, par des méthodes de convergence faible des processus. Nous montrons que la vitesse de convergence est déterminé par la partie martingale du processus. Par conséquent, la vitesse de convergence augmente lorsqu'on simule des variables d'état sous la transformation qui élimine les intégrales stochastiques. La loi asymptotique avec transformation est non-centrée et nous obtenons des expressions de l'erreur d'approximation par calcul de l'espérance de la loi limite. D'autre part, la loi limite sans transformation est centrée et la partie martingale qui détermine la vitesse de convergence est en limite equivalente à une variable aléatoire centrée qui est indépendante des innovations initiales. Par conséquent, l'ordre de convergence de l'espérance de l'erreur est d'ordre supérieur à la convérgence des variables d'état dans ce cas. Nous dérivons des expressions explicites de l'espérance de l'erreur d'approximation comme espérance des variables aléatoires simulables. Ensuite, nous dérivons des lois asymptotiques pour la procédure jointe: solution numérique de l'équation différentielle stochastique et simulation de Monte Carlo pour le calcul des espérances conditionelles. La vitesse de convergence dans ce cas est la même pour les deux methodes. Les expressions des lois limites montrent qu'il existe un biais de deuxième ordre qui correspond à l'espérance de l'erreur d'approximation. Comme nous avons déterminé cette erreur comme espérance des variables aléatoires simulables, nous pouvons alors dérivé des estimateurs pour l'espérance d'une fonction de la valeur terminale d'une equation différentielle stochastique corrigée du biais de deuxième ordre. Finalement, nous dérivons des estimateurs pour les cas où la valeur initiale de variables d'état doit aussi être approximée. Nous montrons que dans ce cas la transformation qui élimine les intégrales stochastiques est de nouveau préférable car l'avantage au niveau vitesse de convergence est préservé comme pour les variables d'état. Nous discutons l'application de ces résultats pour des estimateurs de suivant l'évolution du marché (market timing).

Contents

Remerciements

Je remercie mon directeur de thèse, René Garica, pour sa patience et pour les heures, innombrables, qu'il a consacrées à mon apprentissage. Je lui dois une grosse partie de mes connaissances et il ne dépend que de moi qu'il en reste encore tant à apprendre. J'aimerais aussi remercié Jerôme Detemple, sans lui cette thèse n'aurait pas vu le jour [1].

J'ai également beaucoup profité de mes collègues durant ces quelques années passées au département de sciences économiques. Merci Lars Vilhuber, Stefan Ambec, Karine Gobert, Patrick González, Guido Kuersteiner et les autres, pour vos conseils informatiques, techniques et théoriques. Et surtout, merci pour ces merveilleux moments.

Enfin, un gros merci à Eglantine, Johann et Vivien et mes parents qui m'ont supporté pendant tous le temps de rédaction.

---

[1] René Garcia et Jerôme Detemple sont coauteur des deux derniers essais, qui correspondent aux chapites 2 et 3 de la thèse.

Chapter 1

Insider Information, Arbitrage and Optimal Portfolio and Consumption Policies

1.1 Introduction and Summary

This article extends the standard continuous time financial market model pioneered by Samuelson (1969) and Merton (1971) to allow for investors with anticipative information. Such investors are called insiders, since they can already tell today whether or not an event unknown to the public will occur at some date in the future. The effects of insiders on the social efficiency of mechanisms for intertemporal risk sharing has been of great concern in the financial market literature. The traditional models used to address this issue are based on strong assumptions about the nature of the insider signal and the preferences of investors. In the classical equilibrium models following Grossman (see the articles in Grossman (1989)) and the market microstructure literature based on the market game model of Kyle (1985) it is assumed that insiders have a constant absolute risk aversion and preferences for terminal wealth only. Insider signals are restricted to the liquidation value of the risky asset. Since prices of risky assets are determined endogenously and therefore depends on information, preferences and beliefs of other market participants such anticipative information might be rare. Furthermore these models do not allow for investment horizons for which the uncertainty about the insider signal is resolved given public information.

In this paper we generalize results of Karatzas and Pikovsky (1996a) and show how in a standard continuous time financial market model, techniques of the theory of enlargements of filtrations can be used to analyze the effects of arbitrary insider signals on dynamic portfolio and consumption policies. We illustrate our approach with an example in which an investor knows already today whether or not the risky asset stops to pay dividends after a certain time. Such signals could not been considered previously. For an insider with constant relative risk aversion we derive an explicit expression for the demand of risky assets which is due to his/her anticipative information. More generally the techniques introduced in this paper enable us to analyze how the valuation of contingent claims and optimal consumption and portfolio policies are affected by the precision of the information about the true state of nature contained in insider signals and can therefore also be used for purposes of dynamic risk management.

Since the seminal paper of Harrison and Kreps (1979), it is well-known that the existence of a probability measure for which prices are martingales is equivalent to the existence of a viable market model, that is a model where there exists an optimal trading strategy for some agent with strictly monotonic, continuous and convex preferences. This is often referred to as the fundamental theorem of asset pricing [1]. As shown by Kreps (1981) viability is a stronger requirement than the

[1]The result of Harrison and Kreps (1979) has been generalized in various directions. See Dalang, Morton and Willinger (1990) , Back and Pliska (1991), Schachermayer (1994) for models

absence of free lunches. If the market model is viable, there cannot exist strategies which provide positive gains from trade with positive probability but no initial investment.

We show that the existence of a viable market model for an investor who already knows today whether or not an event unknown to the public will occur, depends crucially on the information about states of nature contained in his/her anticipative information. We prove that for investment horizons which do not end before the first moment in time the informational advantage of an insider has disappeared, there exist always free lunches with vanishing risk but no arbitrage opportunities [2] whenever the insider's anticipative information does not contain zero probability events (atomic insider information). On the other hand if the insider information is so informative about the state of nature that it contains events which are not believed to occur given public information (non-atomic insider information), a viable market model exists only if the investment horizon ends before the uncertainty about the insiders information is resolved. This implies that no viable market model exists for insider whose anticipative information

with discrete trading, Delbaen and Schachermayer (1994) respectively (1998) for general semimartingale models and continuous trading and Jouini and Kallal (1995a) and (1995b), Pham and Touzi (1996) and Wang (1998) for models with constraints on trading strategies. Dybvig and Huang (1988) and Delbaen, Monat, Schachermayer, Stricker and Schweizer (1997) have shown that that the absence of arbitrage can be enforced if gains from trade are bounded in the $\mathbf{L}^p$ norm.

[2]As shown by Delbaen and Schachermayer (1995) there is a difference between the absence of free lunches with vanishing risk and the absence of arbitrage. An investor has free lunches with vanishing risk whenever there exists a sequence of portfolio policies with associated discounted gains from trade bounded from below, such that discounted terminal wealth associated with the sequence of portfolio policies does not converge in probability to the initial wealth. In contrast a portfolio policy is an arbitrage if associated gains form trade a bounded from below with probability one and discounted wealth is bigger than initial wealth with positive probability. See section 1 for precise definitions.

is generated by signals with continuous distributions when the investment horizon is unrestricted. We show that this problem can be avoided if we add independent noise to the insider signal. In this case we can always derive optimal portfolio and consumption strategies. But since in this case uncertainty about the true signal is never resolved it does not capture an important feature of insider information in financial markets and resembles more market models with differential information [3]. Clearly insiders with imperfect anticipative information exist in financial markets, but it seems less realistic that uncertainty about their private signal is never resolved given public information. If we add noise to the true signal we cannot analyze the effects of anticipative information which will be also known by non-insiders at some future point in time. The introduction of independent noise allows just do model noisy insider information such that the information advantage never disappears.

An extension of the fundamental theorem of asset pricing to differential information has previously been considered by Duffie and Huang (1986). They show that in a fully revealing rational expectation equilibrium where better informed agents do not have free lunches, insiders must fully reveal their information to guarantee the existence of a viable market model. Our result shows that their assumption about the existence of absolutely continuous local martingale measures for better informed agents holds true only if the difference of agents' flows of information is atomic when the information advantage can disappear.

We show that the absolutely continuous local martingale measure of an insider and of a non-insider are identical if restricted to public information. Consequently contingent claims whose payoff is measurable with respect to public information must also be valued identically. In contrast, since we show that contingent claims for insiders with non-atomic anticipative information have no value, option pricing

[3] This result also explains how Elliott, Geman and Korkie (1997) prevent the existence of arbitrage opportunities for insiders by introducing insider information on incomplete information.

by arbitrage is invariant to atomic but not to non-atomic anticipative information.

To prove the existence of an optimal trading strategy and therefore the existence of a viable market model for an investor with insider information poses additional difficulties. If we want to solve Merton's consumption-investment problem for such an investor, we face the problem that portfolio policies may depend on anticipative information. Consequently, if we allow for general continuous time investment strategies these may not be adapted to the filtration generated by returns of risky assets and therefore gains from trade can no longer be defined as Itô integrals [4]. Therefore, if we want to allow for general trading strategies we first have to find the representation of the processes relevant for the investment and consumption decision with respect to the insider's enlarged flow of information. If such a representation exists, we can allow for general portfolio policies adapted to the insider's filtration and answer questions about the existence of free lunches simply by checking whether or not local martingale measures exist on the enlarged flow of information.

Arbitrage opportunities for non-atomic information are such that associated gains from trade can replicate any desired wealth process with zero cost. Therefore, an insider who has non-atomic anticipative information does not face any budget constraint and consequently will attain infinite expected utility[5]. Consequently no viable model can exist when the insiders investment horizon ends after his/her private information is completely revealed and insider information is

[4] One way to deal with this problem would be to use results from anticipative stochastic calculus and to define gains from trade as Skorohod integrals. The drawback of such an approach is that the semi-martingale property of wealth which is linked to the absence of free lunches is completely lost. In discrete time models the definition of gains from trade with anticipative strategies is still possible but the resulting process will not be a martingale.

[5] This result for non-atomic insider information therefore explains why Karatzas and Pikovsky (1996a) find infinite additional logarithmic utility from final wealth in a Gaussian model for signals corresponding to final states or final prices without noise.

non-atomic.

The solutions of Merton's consumption-investment problem for insider information previously presented in Karatzas and Pikovsky (1996a), Eliott, Geman and Korkie (1997) as well as Amendinger, Imkeller and Schweizer (1998) are based on von Neumann-Morgenstern preferences for terminal wealth with logarithmic utility. Since such preferences lead to myopic portfolio policies which can be obtained by a sequence of one-period optimization problems, they avoid questions about the existence of hedging portfolio policies which finance optimal cumulative consumption adapted to insider information. For logarithmic preferences we do not have to be concerned about the perfect replicability of the cumulative consumption. Since perfect hedging for insider information given as an initial enlargement of a Brownian filtration is still possible, we are able to solve Merton's consumption-investment problem for more general utility functions, independently of whether preferences are defined over terminal wealth, consumption or both.

All the results in this paper are based on partial equilibrium considerations. We focus on the consumer's problem who takes a certain flow of information as given. In this article we are not asking if this information is actually implementable in equilibrium. Clearly the insider's optimal strategies for atomic anticipative information can always be supported in a fully revealing rational expectation equilibrium. But since we show that optimal consumption and portfolio policies do not reveal all anticipative information, a Walrasian auctioneer will not learn all the insider information from the individual demands. Consequently, such an equilibrium will not exist without other channels of information transmission. Such issues will be discussed in more detail in Rindisbacher (1998).

Similarly our results have important consequences for many market microstructure models. Such models are generally principal-agent models where the principal is the market maker and the agent is the insider. For the regulation of insider trading these models imply that the allocative efficiency is improved whenever a mechanism can be found by which insiders reveal more information as in any model with adverse selection. Our results about the non-existence of a viable market model for non-atomic insider information imply that this is not necessarily the case. For such signals full revelation of insider information will lead to a breakdown of markets for intertemporal risk sharing. In a companion paper (Rindisbacher (1998)) we show that this effect more generally known as the Hirshleifer (1971) effect in the insurance literature may play an important role if we want to analyze the social costs of insider trading [6].

The paper is organized as follows. In section 1.2 we introduce the models for public information and derive the representation of the price, endowment and wealth processes for the insider information. In section 1.3 we analyze whether or not anticipative information allows for arbitrage opportunities. Then in section 1.4 we consider the pricing of contingent claims for insiders. The results in section 1.4 are necessary for the existence of a solution of Merton's consumption-investment problem for an insider. Explicit expression for optimal consumption and portfolio policies for insiders are presented in section 1.5. In section 1.6 the results are illustrated with two examples.

In Appendix 1.7 we show in detail how the model relevant for insiders can be obtained from the "Girsanov approach" to initial enlargements of filtrations. In Appendix 1.8 we show how random variables measurable with respect to an enlarged filtration can be represented as product-measurable functions of signals and states. The two appendices contain all results necessary to solve Merton's consumption-investment problem for an insider. Appendix 1.9 contains some definitions from Malliavin calculus. Finally we present the proofs of our results in

[6] We argue that this effect is the source for the findings of Back (1993), that in the presence of asymmetric information it may be impossible to price options by arbitrage and is not captured in principal-agent models for insider trading.

Appendix 1.10.

1.2 A Model for Public and Insider Information

A financial market model can be characterized by $(\Omega, \mathcal{F}_1, \mathbb{F}, \mathbf{P}, \mathcal{C}, \succeq)$ where $(\Omega, \mathcal{F}_1, \mathbb{F}, \mathbf{P})$ corresponds to a stochastic basis consisting of a state space Ω, possible information $\mathcal{F}$, flow of information $\mathbb{F}$ and beliefs $\mathbf{P}$. The space $\mathcal{C}$ denotes the consumption space on which a preference ordering $\succeq$ is defined. In this section we first introduce the model for public information and then show how from this we obtain the model for the insider by initial enlargements of filtrations.

1.2.1 The Model for Public Information

We consider a frictionless market where each investor has the choice between d dividend paying risky assets and one asset without risk. Possible states of nature are given as points in a d-dimensional Wiener space $\Omega = C^0([0,1];\mathbb{R}^d)$. Possible information is given by the Borel σ- field $\mathcal{F}_1$ on Ω. Investors' beliefs are homogeneous and given by the standard Wiener measure $\mathbf{P}$, the measure for which observed states given as trajectories of the coordinate process $W = (\omega(t), t \in [0,1])$ on Ω correspond to a d-dimensional Wiener process. The flow of information $\mathbb{F} := (\mathcal{F}_t)_{t\in[0,1]}$ available to all investors is defined as the $\mathbf{P}$-completion of the Wiener filtration, the natural filtration of the coordinate process W.

Given the topology of the state space Ω, all processes on the stochastic basis $(\Omega, \mathcal{F}_1, \mathbb{F}, \mathbf{P})$ relevant for investors' decisions are given as Brownian functionals adapted to the flow of information which is given by the Wiener filtration $\mathbb{F}$.

The only risk free asset B (i.e. the predictable asset of bounded total variation) is defined as an exponential of a bounded adapted interest rate process $r = (r_t, t \in$

$[0,1])$ by

$$B_t := \exp(\int_0^t r_s ds) \tag{1.1}$$

whereas the $(\mathbf{P}, \mathbb{F})$- Doob-Meyer decomposition of the d-dimensional vector of risky assets $P = ((P_t^j)_{j=1,\dots,d}; t \in [0,1])$ is given by

$$P_t + \int_0^t D_s ds = P_0 + \int_0^t diag[P_s^j][b_s ds + \sigma_s d\omega(s)] \tag{1.2}$$

where $D = ((D^j)_{j=1,\dots,d}; t \in [0,1])$ denotes the dividend process. This process is assumed to be exogenous and given by

$$D_t = D_0 + \int_0^t diag[D_s^j]\mu^D(s, D_s)ds + \int_0^t diag[D_s^j]\gamma^D(s, D_s)d\omega(s). \tag{1.3}$$

The other exogenously given process in the economy is the endowment rate process which satisfies the following stochastic differential equation

$$e_t = e_0 + \int_0^t \mu^e(s, e_s)ds + \int_0^t (\gamma^e(s, e_s))^* d\omega(s). \tag{1.4}$$

It follows that if the initial dividend D_0 is positive dividends will be always positive whereas endowments are allowed to be negative. For both equations we assume that coefficients satisfy global Lipschitz conditions that guarantee existence and linear growth conditions that are sufficient to get unique solutions [7]. Furthermore we assume that coefficients are differentiable to an appropriate degree.

The coefficients (r, b, σ) of the model are assumed to be bounded and $\mathcal{F}_t$-adapted. The volatility coefficient σ is assumed to be positive definite $\mathbf{P} \otimes \boldsymbol{\lambda}$ a.e. on the product space $L^2([0,1] \times \Omega)$ of random functions, where $\boldsymbol{\lambda}$ corresponds to the Lesbegue measure. Furthermore we assume that $b_t^i \in \mathbb{L}^{1,2}(\mathbb{R}^q)$ and $\sigma_t^{i,j} \in \mathbb{L}^{1,2}(\mathbb{R}^q)$ for all $i, j \in \{1, \dots, d\}$ where $\mathbb{L}^{1,2}$ corresponds to the domain of the Skorohod integral which is defined in Appendix 1.9[8].

[7]see Nualart (1995) p.99 for example

[8]The assumptions made about price coefficients are stronger than required to solve the

An admissible trading strategy for risky assets is a d-dimensional adapted vector process $\pi = ((\pi^j)_{j=1,\ldots,d}; t \in [0,1])$ such that $\pi \in \mathbb{L}_a^{1,2}(\mathbb{R}^q)$ where the subscript denotes the restriction to adapted processes in $\mathbb{L}^{1,2}$. To exclude free lunches from doubling strategies for an investment horizon $T \in [0,1]$ we assume that optimal strategies are tame[9], meaning that the corresponding wealth $X^{\pi,c}$ at any moment in time $t \in [0,T]$ must be bounded from below

$$\frac{X_t^{\pi,c}}{B_t e_0} \geq -K \tag{1.5}$$

P- a.s. for some $K > 0$.

The consumption space is given by $\mathcal{C} = \mathbb{L}_a^{1,2}(\mathbb{R}^q)$. We assume that consumption is absolutely continuous with respect to the Lesbegue measure and can therefore be written as $C_t = \int_0^T c_s ds$. The wealth process $X^{\pi,c} = (X_t^{\pi,c}; t \in [0,1])$ satisfies the following stochastic differential equation.

$$X_t^{\pi,c} = e_0 + \int_0^t X_s^{\pi,c} r_s ds + \int_0^t (\pi_s)^* \sigma_s[\theta_s ds + d\omega(s)] - \int_0^t (c_s - e_s) ds, \tag{1.6}$$

where $\theta = (\theta_t; t \in [0,1])$ denotes the market price of risk or conditional Sharpe ratio, defined by $\sigma_t \theta_t := b_t - 1_d r_t$.

Final wealth must be non-negative at the end of the investment horizon for strategies to be admissible

$$\frac{X_T^{\pi,c}}{B_T e_0} \geq 0. \tag{1.7}$$

---

consumption-investment problem for Brownian information $\mathbb{F}$, but enable us to find explicit expressions for optimal portfolio policies from Malliavin's calculus. Furthermore, the techniques presented in this paper to get explicit hedging strategies for contingent claims measurable with respect to insider information apply to claims H for which Malliavin derivatives exist, that is $H \in \mathbb{D}^{1,1}(\mathbb{R}^q)$, where $\mathbb{D}^{1,1}$ denotes the domain of the Malliavin derivative defined in Appendix 1.9.

[9]Tameness of portfolio policies is necessary to exclude doubling strategies (see Dybvig and Huang (1988))

Preferences $\succeq$ are of the von Neumann-Morgenstern type with an additive state independent utility function

$$U(T, c, \mathbf{P}, \mathbb{F}) := \mathbf{E}^{\mathbf{P}}[\int_0^T u(s, c_s)ds|\mathcal{F}_0], \tag{1.8}$$

where the utility function $u \in C^{1,2}([0,T] \times \mathbb{R}_+; \mathbb{R})$ is strictly increasing and strictly concave. It also satisfies the Inada conditions $\lim_{c\to 0} \partial_2 u(t,c) = +\infty$ and $\lim_{c\to+\infty} \partial_2 u(t,c) = 0$. The corresponding absolute risk aversion is given by $A(t,c) := -\partial_2 \log \partial_2 u(t,c)$. The inverse of marginal utility I is defined by $\partial_2 u(t, I(t,y)) = y$. We require that investment-consumption strategies (π, c) satisfy the condition

$$\mathbf{E}^{\mathbf{P}}[\int_0^T u^-(t, c_t)dt|\mathcal{F}_0] < \infty \tag{1.9}$$

where $u^- := -min(0, u)$. Without this technical condition the consumption-investment problem is not well posed [10]. If (π, c) satisfies the budget constraints (1.6) as well as conditions (1.5), (1.7) and (1.9) we call $\mathcal{F}_t$-adapted (π, c) admissible and write $(\pi, c) \in \mathcal{A}(\mathbf{P}, \mathbb{F}, e)$ meaning that admissibility holds for beliefs $\mathbf{P}$, flow of information $\mathbb{F}$ and endowment process e.

Finally we introduce the following deflator process $S = (S_t; t \in [0,1])$ where

$$S_t := -\int_0^t r_s ds - \int_0^t (\theta_s)^* d\omega(s). \tag{1.10}$$

The corresponding state price density process $\mathcal{E}(S) = (\mathcal{E}(S)_t; t \in [0,1])$ can then be written as a stochastic exponential of the deflator process, that is as the unique solution of the following stochastic differential equation

$$\mathcal{E}(S)_t = 1 + \int_0^t \mathcal{E}(S)_v dS_v.$$

[10] see Karatzas, Lehocky, Sethi and Shreve (1986) for a discussion.

1.2.2 The Model for Insider Information

To analyze the effects of anticipative information on consumption and portfolio policies, we propose a model for the insider given by $(\Omega, \mathcal{F}_1, \mathbb{G}, \mathbf{P}, \mathcal{C}, \succeq)$, where his/her flow of information $\mathbb{G} := (\mathcal{G}_t)_{t\in[0,1]}$ is obtained by an initial enlargement of filtration

$$\mathcal{G}_t := \bigcap_{\epsilon>0}[\mathcal{F}_{t+\epsilon} \bigvee \sigma(G)], \quad t \in [0,1], \tag{1.11}$$

with G a q-dimensional random vector with law $\mathbf{P}_G$. The random vector G corresponds to the signal which, at the beginning of the investment horizon, is only known by the insider. Signals previously considered in the literature correspond to terminal states of nature (Karatzas and Pikovsky (1996a) and Elliott, Geman and Korkie (1997)), terminal values of risky assets (see the articles in Grossman (1990), Kyle (1989) and Karatzas and Pikovsky (1996a)) or indicators of these variables (Karatzas and Pikovsky (1996a)). In what follows we allow for any kind of $\mathcal{F}_1$- measurable signals.

To the signal G we can associate the resolution time $T_G := inf\{t \in [0,1] : \mathbf{E}[\mathbf{1}_E|\mathcal{F}_t] = \mathbf{1}_E \forall E \in \mathcal{G}_t\}$, meaning that the stopping time T_G designates the first moment in time at which the information advantage about the signal G that was known at the beginning ($t = 0$) has disappeared. Clearly, the signal G is $\mathcal{F}_{T_G}$ measurable.

In Appendix 1.7 we show in detail how we obtain the representation of processes on the stochastic basis $(\Omega, \mathcal{F}_1, \mathbb{G}, \mathbf{P})$ relevant for the insider's investment decision. The assumptions used to derive these decompositions and the results in the following sections are as follows

Assumption 1 *(**"Condition A" Jacod (1980)**) There exists a common measure $\boldsymbol{\nu}$ on the Borel field $\mathcal{B}_{R^q}$ such that $\mathbf{P}_t^\omega \ll \boldsymbol{\nu}$ for all $t \in [\![0, T_G[\![$ where $\mathbf{P}_t^\omega$ corresponds to the conditional law of G given the initial filtration $\mathcal{F}_t$.*

This assumption basically guarantees the stability of semi-martingales under initial enlargements of filtrations ("Hypothèse H" in Jacod (1980)) on the stochastic interval $[\![0, T_G[\![$. It is necessary since if the insider has a private signal such that prices with respect to the enlarged flow of information are not any longer semi-martingales no viable market model exists. As shown by Delbaen and Schachermayer (1994) the semi-martingale property is necessary for the absence of free lunches. Our assumptions on the coefficients (r, b, σ) guarantee that after resolution time all processes relevant for the insider's investment decisions are semi-martingale since on $[\![T_G, 1]\!]$ as we will see below the insider's model is identical to the model for public information.

If "condition A" is satisfied it follows from the Radon-Nikodym theorem that $\mathbf{P}_t^\omega(dz) = p(\omega, t, z)\boldsymbol{\nu}(dz)$ for all $t \in [\![0, T_G[\![$. For $t = 0$ we get the unconditional density of the signal as Radon-Nikodym derivative $\mathbf{P}_0^\omega(dz) = q(z)\boldsymbol{\nu}(dz)$, that is $q(z) := p(\omega, 0, z)$. If such a measure $\boldsymbol{\nu}$ exists we can with out loss of generality assume that it corresponds to the unconditional law of the signal $\boldsymbol{\nu} = \mathbf{P}_G$. The following assumption from Imkeller (1996) guarantees that the contemporaneous Malliavin derivative used to represent the conditional density process as a non-negative martingale is well defined.

Assumption 2 *The conditional density of the signal $p(\omega, t.z)$ is such that (i) $p(\omega, s, z) \in \mathbb{L}^{1,2}(\mathbb{R}^d)$ as well as (ii) the mapping $r \mapsto \mathcal{D}_r^j p(\omega, s, z)$ is left-continuous in $L^1(\Omega)$ at $s \in [0, t]$ and $z \in \mathbb{R}^q$, for all $j \in \{1, \cdots, d\}$, where $\mathcal{D}_t^j p(\omega, t, z)$ denotes the Malliavin derivative of the conditional density.*

This assumption will be used to derive the drift of the processes relevant for the insider's investment decision. Given the conditional density of the signal it allows to define the process $\alpha_t^z(\omega) := \lim_{s \uparrow t} \frac{\mathcal{D}_s p(\omega, t, z)}{p(\omega, t, z)}$. As we explain in Appendix 1.7 this process is used to derive the $(\mathbf{P}, \mathbb{G})$- Brownian motion W^G given in (1.114). The

Brownian motion W^G on the insider information is such that $dW_t^G = d\omega(t) - \alpha_t^G dt$ for $t \in [\![0, T_G[\![$ and $dW_t^G = d\omega(t)$ for $t \in [\![T_G, 1]\!]$.

Finally the next assumption is necessary for $(\mathbf{P}, \mathbb{G})$- semi-martingales to exist up to the moment at which the information advantage of the insider has disappeared.

Assumption 3 *Signals G are such that*

$$\int_0^{T_G} \left| \left[\frac{\mathcal{D}_t^j p(\omega, t, z)}{p(\omega, t, z)} \right]_{|z=G(\omega)} \right| ds < \infty \tag{1.12}$$

$\mathbf{P}$-*a.s. for all $j \in \{1, \cdots, d\}$, where $\mathcal{D}_t^j p(\omega, t, z)$ denotes the Malliavin derivative of the conditional density.*

Without this assumption price processes for insider information are not semi-martingales on $[\![0, T_G[\![$. Consequently the insider will have free lunches and no viable model for such investors exists.

The following theorem uses the results derived in Appendix 1.7 and provides the market model from the point of view of the insider.

Theorem 1 *For insider signals G such that assumption 1, 2 and 3 are satisfied, the representation of price, dividend rate and endowment processes on the stochastic basis relevant for the insider $(\Omega, \mathcal{F}_1, \mathbb{G}, \mathbf{P})$ is for $t \in [0, 1]$ as follows.*

Risk free assets are given by:

$$B_t = \exp\left(\int_0^t r_s ds\right). \tag{1.13}$$

The prices of risky assets P satisfy:

$$P_t + \int_0^t D_s ds = P_0 + \int_0^t diag[P_s^j][b_s + \sigma_s dW_s^G] + \int_0^{t \wedge T_G} \sigma_s \alpha_s^G ds, \tag{1.14}$$

where the dividend process D is given by

$$D_t = D_0 + \int_0^t diag[D_s^j][\mu^D(s, D_s)ds + \gamma^D(s, D_s)dW_s^G] + \int_0^{t\wedge T_G} \gamma^D(s, D_s)\alpha_s^G ds. \quad (1.15)$$

The endowment rate satisfies the following stochastic differential equation

$$e_t = e_0 + \int_0^t [\mu^e(s, D_s)ds + (\gamma^e(s, e_s))^* dW_s^G] + \int_0^{t\wedge T_G} \gamma^e(s, e_s)^* \alpha_s^G ds. \quad (1.16)$$

For strategies $(\pi, c) \in \mathcal{A}(\mathbb{G}, \mathbf{P}, e)$ *such that* $\int_0^{T_G} |\pi_s^* \sigma_s \alpha_s^G| ds < +\infty$ $\mathbf{P}$- *a.s. we get the following discounted wealth process*

$$\frac{X_t^{\pi,c}}{B_t} = e_0 + \int_0^t \frac{(\pi_s)^* \sigma_s}{B_s}[(\theta_s + dW_s^G] - \int_0^t \frac{(c_s - e_s)}{B_s} ds + \int_0^{t\wedge T_G} \frac{(\pi_s)^* \sigma_s \alpha_s^G}{B_s} ds, \quad (1.17)$$

We see that anticipative information affects the insider's view about the individual market price of risk, the appreciation rate of the dividend and price processes of the risky assets as well as the growth rate of the endowments in a way which depends on the "contemporaneous elasticity with respect to changes of the state of nature" (i.e. logarithmic Malliavin derivative) of the signal's conditional density.

The anticipative information does not change the quadratic variation of the processes relevant for the portfolio choice. If we compare the price processes of insiders and non-insiders we see that they agree on the volatility of the asset prices but not on their appreciation rates. A priori this seems to be surprising but it simply stems from the fact that for price processes given as semi-martingales the quadratic variation is locally already known by non-insiders (i.e. $\mathbb{F}$- predictable). As a consequence additional information cannot locally reduce the conditional volatility of the process. Exactly for the same reason the difference in information lets the price process of a risk free (i.e. predictable) asset unchanged.

Since the initial information of an insider $\mathcal{G}_0 = \sigma(G)$ is, in contrast to the public information $\mathcal{F}_0 = \{\Omega, \emptyset\}$ non trivial, his/her beliefs are not any longer given by $\mathbf{P}$ but by the conditional probabilities $\mathbf{P}(\cdot|\mathcal{G}_0)$. In Appendix 1.7 we show that $\mathbf{P}(\cdot|\mathcal{G}_0) = \mathbf{P}^z_{|z=G}$ where $\mathbf{P}^z$ denotes the conditional Wiener measure which concentrates its probability mass on $\{G = z\}$. The conditional Wiener measure is obtained from the joint law of the signal and states of nature, defined on the product space $\Omega \times \mathbb{R}^d$. Therefore, states relevant for an insider can be described in form of pairs (ω, G) and beliefs take all its probability mass along the diagonal of the product state space. In what follows this fact will play an important role. Because of a decoupling property of the insider's local martingale measure, we can derive optimal hedging policies first by conditioning on the realization of the signal. If evaluated at the true signal, these policies will be shown to be optimal for an insider.

We have already seen that anticipative information does not affect the volatility coefficient of risky assets. The following corollary gives necessary and sufficient conditions which guarantee that the drift coefficients of the processes relevant for the portfolio choice remain unchanged.

Corollary 2 *Under the conditions of theorem 1 an insider's decompositions of price, dividend and endowment processes on* $(\Omega, \mathcal{F}_1, \mathbb{F}, \mathbf{P})$ *and* $(\Omega, \mathcal{F}_1, \mathbb{G}, \mathbf{P})$ *are the same if and only if for all* $t \in [\![0, T_G[\![$ *the information generated by the signal is independent from the common available information.*

$$\mathcal{F}_t \perp \sigma(G) \tag{1.18}$$

It follows that information which is independent of the flow of common information for given beliefs is irrelevant for the insider's investment decision since it does not change his/her wealth process. Then if the events revealed by his/her side information do not increase knowledge about initial pay-off relevant events, the observation of such a signal does not reduce uncertainty concerning states of nature relevant for his/her consumption and investment decision. Such information is therefore simply not taken into account. Insider information only affects optimal strategies if it helps to get a finer flow of information concerning events the he cares about. We will call independent signals redundant, meaning that they are irrelevant for investors' decisions. The redundancy of independent signals implies that consumption and investment strategies are unaffected by idiosyncratic shocks which are independent from the common flow of information. Consequently, such shocks will have no impact on equilibrium prices [11].

1.3 Insider Information and Arbitrage

Based on the insider's model obtained in theorem 1 we now investigate whether or not an insider has necessarily free lunches. We first define two notions of free lunches previously considered in the literature: the concept of free lunches with vanishing risk and the concept of arbitrage opportunities. We also introduce the concept of conditional arbitrage which seems more appropriate when initial information is non-trivial as for an insider. Then we explain how we distinguish between atomic and non-atomic insider information. The next section shows that dependent on the investment horizon these distinctions matter for an insider. Our results illustrate that it is the precision of the anticipative information about the state of nature which provides an insider with arbitrage opportunities and not the fact that he has more information than a non-insider alone.

[11] See Rindisbacher (1998) for more on this point.

1.3.1 Definitions

If assumption 3 is not satisfied, prices of risky assets are not any longer semimartingales and as a consequence it follows from results in Delbaen and Schachermayer (1994) that the insider has free luches with vanishing risk, that is their NFLVR condition is not satisfied. In our context we can state the NFLVR condition for an investment horizon T as follows.

Definition 3 *The price process of risky assets satisfies the NFLVR condition if for all sequences of portfolio strategies $(\pi^n)_{n\in N}$ with associated wealth process such that*

$$\frac{X_T^{\pi^n,e}}{e_0 B_T} < +\infty \tag{1.19}$$

$\mathbf{P}$-*a.s. and positive sequences $(\delta_n)_{n\in N}$ converging to* 1 *such that*

$$\frac{X_t^{\pi^n,e}}{e_0 B_t} > \delta_n, \tag{1.20}$$

$\mathbf{P}$- *a.s. for all $t \in [0,T]$, we have that*

$$\mathbf{P} - \lim_{n\to+\infty} \frac{X_T^{\pi^n,e}}{e_0 B_T} = 1. \tag{1.21}$$

Condition (1.19) excludes portfolio policies for which wealth is not defined and the sequence δ^n in (1.20) guarantees that gains from trade are bounded from below and therefore excludes doubling strategies. As shown in Delbaen and Schachermayer (1994) and also in Back and Pliska (1991) this condition is more restrictive than the no arbitrage condition. Insiders do not necessarily have arbitrage opportunities if they have free lunches with vanishing risk. As we will show below the distinction between free lunches with vanishing risk and arbitrage opportunities is important for insider information.

In our model we define an arbitrage opportunity as follows

Definition 4 *A tame portfolio policy π with associated wealth process $X^{\pi,e}$ such that*

$$\mathbf{P}(\frac{X_T^{\pi,e}}{e_0 B_T} \geq 1) = 1 \tag{1.22}$$

and

$$\mathbf{P}(\frac{X_T^{\pi,e}}{e_0 B_T} > 1) > 0 \tag{1.23}$$

for some $T \in [0,1]$ is called an arbitrage.

This basically states that if gains from trade are almost surely believed to be non-negative but not strictly positive there is no arbitrage. Since convergence in probability does not imply almost sure convergence a free lunch is not necessarily an arbitrage. Behind both definitions is the idea that a reasonable model of the financial market should not allow investors to make positive gains from trade out of nothing.

The definitions above previously considered in the literature are based on the assumption that the information of the investor at the beginning of his/her investment horizon corresponds to the trivial information set $(\Omega, \emptyset)$. Since we want to analyze the flow of information of an insider who has anticipative information given by $\sigma(G)$ already at the initial date, we need a notion of conditional arbitrage.

Definition 5 *A conditionally tame portfolio policy (π such that for all $t \in [0,1]$ we $\mathbf{P}$- a.s. have that $\mathbf{P}(\frac{X_t^{\pi,c}}{B_t e_0} \geq -K|\mathcal{G}_0) = 1$) with associated wealth process $X^{\pi,e}$ such that $\mathbf{P}$- a.s.*

$$\mathbf{P}(\frac{X_T^{\pi,e}}{e_0 B_T} \geq 1|\mathcal{G}_0) = 1 \tag{1.24}$$

and for some $E \in \mathcal{G}_0$ where $\mathbf{P}(E) > 0$

$$\mathbf{P}(\frac{X_T^{\pi,e}}{e_0 B_T} > 1|\mathcal{G}_0) > 0 \tag{1.25}$$

for some $T \in [0,1]$ is called a conditional arbitrage.

Therefore, for an absence of arbitrage it is necessary that there is no arbitrage given any event known at the beginning of the investment horizon, whereas for a conditional arbitrage to exist it is sufficient that gains from trade are believed to be positive given the initial information. This shows that the absence of arbitrage is a stronger assumption than the absence of conditional arbitrage.

Proposition 6 *For a given flow of information an investor who has a conditional arbitrage has necessarily an arbitrage, whereas an investor who has an arbitrage may not have a conditional arbitrage.*

This proposition does not compare arbitrage opportunities for insiders and non-insiders since it assumes that portfolio policies are adapted to the same flow of information. It implies that it is sufficient to prove the absence of an arbitrage for the enlarged flow of information to guarantee the existence of a viable model for the insider.

As shown by Harrison and Kreps (1979) it is necessary for the existence of a pricing kernel that prices any asset that there is no arbitrage. In what follows we analyze how this result depends on the flow of information available to investors. We will show how the informational content of the insider signal determines whether or not an insider has arbitrage opportunities. It will be important to distinguish between atomic and non-atomic anticipative information

Definition 7 *An investor has atomic insider information if there exists an event $E \in \mathcal{G}_0$ such that $\mathbf{P}(E) > 0$ and for all $F \in \sigma(G)$ we have that $F \subset E$ implies either $F = E$ or $F = \emptyset$. If there does not exist such an event E we say that the insider information is non-atomic.*

Clearly if there exists a countable partition of the state space which generates the anticipative information $\sigma(G)$ then there exists a countable number of atoms.

In contrast an investor who has non-atomic insider information knows whether or not an event has occurred out of an uncountable number of events. In this sense atomic insider information is less informative than non-atomic insider information. Signals with discrete distributions generate atomic insider information, whereas signals with continuous distributions generate non-atomic insider information. Non-atomic insider information in contrast to atomic insider information contains events whose occurrence is not believed given public information. This fact will be crucial for whether or not an insider has an arbitrage opportunity.

1.3.2 Main Result

We now show that the distinction between different concepts of free lunches for insiders matters. The main result of this section shows how the existence of arbitrage opportunities for insiders does depend on the complexity of their anticipative information.

Theorem 8 *Under the assumptions of theorem 1 we have for insider information which is not independent of the common available flow of information and non-trivial ($\sigma(G) \neq \{\Omega, \emptyset\}$) that*

1. *Any investor who knows the information revealed by G and who has an investment horizon such that $T \in [\![T_G, 1]\!]$ has necessarily free lunches with vanishing risk.*

2. *Any investor who knows the information revealed by G has no arbitrage if and only if either $T \in [\![0, T_G[\![$ or $T \in [\![T_G, 1]\!]$ and insider information $\mathbb{G}$ is atomic.*

As it follows from the proof, the results of the theorem are basically a consequence of the fact that the unconditional law of the signal cannot any longer be absolutely continuous with respect to the conditional law of the signal when the uncertainty about the insider information is resolved. This implies that the state price density process which depends on the conditional and unconditional law of the signal is not anymore strictly positive after the realization of the signal is known to the public. It follows that there can exist at most an absolutely continuous martingale measure. On the other hand if there is no free lunch with vanishing risk the density process of the local martingale measure must be strictly positive [12] and therefore the unconditional law of the signal would be absolutely continuous with respect to the conditional law after resolution time. Since this is only possible if the signal is constant there must exist free lunches with vanishing risk for such investment horizons. Furthermore since if there does not exist any atoms, the conditional and unconditional laws of the insider signal are even mutually singular, not even an absolutely continuous local martingale measure will exist when the investment horizon does not end before his/her information advantage has disappeared. In contrast, we show that the equivalence of the conditional and unconditional law of the signal before the information advantage is lost, also excludes arbitrage opportunities if the investment horizon ends before the insider information is fully known by the other investors.

If an insider is only allowed to trade before the resolution time, he cannot attain arbitrage opportunities. Similarly, if he is not allowed to trade before the uncertainty about the information contained in his signal is resolved he has no anticipative information and cannot obtain an arbitrage. We show in the proof that a mean-variance demand for risky assets at the resolution time and zero elsewhere is an arbitrage if insider information is non-atomic [13]. Therefore, if an insider is able

[12] See also Delbaen and Schachermayer (1995).

[13] The existence of an absolutely continuous martingale measure depends critically on the behavior of $\lim_{T \mapsto T_G} \int_{]t,T]} ||\theta_s + \alpha_s^z||^2 ds$ some $t \in [0, T_G[$ under the conditional Wiener measure $\mathbf{P}^z$ as it follows from the arguments in Delbaen and Schachermayer (1995) and Levental and Skorohod (1995) . If the signal is discrete this expression reaches smoothly $+\infty$ whereas

to realize arbitrage opportunities he can do this with strategies which trade just immediately before his/her information advantage has disappeared. This implies that any effective regulation of insider trading must prevent insiders from trade contingent upon anticipative information on a time span that includes resolution time.

The fact that for insiders without restrictions on the investment horizon only absolutely continuous local martingale measures exist, seems problematic from the point of view of financial innovation. In this case there exist binary options paying zero dollars or one dollar contingent upon events of the form $\{G = z\}$ some $z \in \mathbb{R}^d$ which are known not to occur by the insider only. Such a claim has no value for insiders but a positive value for non-insiders. There will be an infinite offer for such claims and consequently infinite profits for insiders issuing them. But to develop such an argument we must take into account that the marketing of those claims will necessarily change the distribution of information in the economy since the claim is written on an event unknown to non-insiders. This will affect the flow of information and therefore the set of local martingale measures, for non-insiders in such a way that these claims may have no value for initial non-insiders as well.

In many market models following the work of Grossman (see the articles in Grossman (1990)) it is assumed that the insider's signal is given by the "true signal" G^0 perturbed by some independent noise Z. The following corollary shows that this prevents arbitrage opportunities for the insider.

---

for non-atomic signals it jumps to $+\infty$. In the proof of their main result they use that the existence of an absolutely but non-equivalent local martingale measure is equivalent to the NFLVR condition under a probability which restricts its positive mass on events on which we have absolute continuity. In our model these events are given by $\{G = z\}$ which are of positive probability only if the signal is discrete. As a consequence a local martingale measure will only exist in this case.

Corollary 9 *Under the assumptions of theorem 1 we have for signals G^0 which are perturbed by independent noise Z*

$$G = G^0 + Z \tag{1.26}$$

that there is no arbitrage opportunity.

As we show in the proof, in the presence of independent noise, the insider's information advantage never disappears. From this point of view the result of the corollary seems surprising. But since the information advantage always exists, the information contained in the true signal G^0 is never fully revealed and as a consequence perturbed insider information is less informative than true anticipative information and does not allow for arbitrage opportunities. This illustrates that it is not the fact that an insider has more information than the non-insider which provides him arbitrage opportunities but the precision of the information about the states of nature contained in his/her signal.

Before we show how theorem 8 explains why Merton's consumption-investment problem for insiders' having non-atomic anticipative information has no solution we prove that claims adapted to insider information can be replicated without tracking error. This result is crucial for the existence of a solution to the portfolio choice problem with "non-myopic" preferences. It guarantees that insiders can finance any consumption policy adapted to their enlarged flow of information.

1.4 Insider Information and Hedging of Claims

In this section we consider the effects of anticipative information on the valuation of contingent claims. Since an insider has an enlarged set of admissible portfolio and consumption strategies he is able to replicate every claim a non-insider can synthetize and therefore his/her willingness to pay for a contingent claim is necessarily bounded from above by the highest price at which a non-insider is willing to buy. We first prove that an investor can replicate any claim measurable with respect to his/her enlarged information. This is the key result to solve Merton's consumption-investment problem in the next section. Then we show that the valuation of contingent claims by arbitrage is invariant with respect to atomic insider information. In contrast the implicit price of contingent claims is zero if anticipative information contains zero probability events given public information.

1.4.1 Claims measurable with respect to Insider Information

Since we want to use the results presented in this section to solve Merton's consumption-investment problem in a non-Markovian market for an insider we have to allow for claims whose pay-off at maturity may be uncertain for non-insiders. From a technical point of view the basic problem is that in contrast to a Wiener filtration it is not clear whether any $(\mathbf{P}, \mathbb{G})$ local martingale can be written as a stochastic integral with respect to the $(\mathbf{P}, \mathbb{G})$- Brownian motion W^G. If $\mathbb{G}$ does not have the predictable representation property, hedging without tracking error will not be possible for contingent claims which are measurable with respect to the enlarged filtration only. As a consequence the martingale techniques used to solve the consumption-investment problem in a non-Markovian market will not work for general convex preferences other than logarithmic utility for final wealth.

Karatzas and Pikovsky (1996b) have shown that for enlargements with random vectors G such that $G = \eta + \int_0^1 g(t)d\omega(t)$ where η is a random vector independent of $\mathcal{F}_1^\omega$ and $g(\cdot)$ is a deterministic matrix function locally bounded in the operator norm and such that $\int_0^1 ||g(t)||^2 dt < \infty$, any local $(\mathbb{P}, \mathbb{G})$ martingale starting at zero can be represented as a stochastic integral with respect to W^G. It follows that $\mathcal{G}_T$- measurable random variables ("contingent claims") that can be written as the sum of a random variable which lies in the first Wiener chaos plus a random variable which is independent of the flow of publicly available information can be perfectly replicated if markets are complete. We show that this remains true for initial enlargements and contingent claims that are sufficiently smooth.

The basic idea exploited to hedge a $\mathcal{G}_T$- measurable contingent claim H uses theorem 22, where we show that claims adapted to insider information can be written as $\mathcal{B}_{R^q} \otimes \mathcal{F}_T$- measurable random function $C^z(\omega)$ given as shown by

$$C^z(\omega) = \mathbf{E}[H|\mathcal{F}_T](\omega) + \mathbf{E}^{\mathbf{P}^z}[\int_T^{T_G} \mathbf{E}[\mathcal{D}_t H|\mathcal{F}_t]\alpha_t^z dt|\mathcal{F}_T](\omega) \tag{1.27}$$

$\mathbf{P}$- a.s.. As a consequence we are able to replicate contingent claims in two steps. First conditional on the event $\{G = z\}$ for some $z \in \mathbb{R}^d$ we get hedging strategies to hedge C^z. The hedging strategies obtained this way are measurable functions of the realization of the signal. If we evaluate these functions at the true signal $z = G$ we get hedging strategies for H since the absolutely continuous local martingale measure of the insider decouples the signal and the states of nature. The associated minimal cost of the replicating strategies provides the implicit price of the claim.

Theorem 10 *If "condition A"and assumption 2 are satisfied then*

1. *For any investment horizon $T \in [\![0, T_G[\![$ a $\mathcal{G}_T$- measurable contingent claim H such that $\mathcal{E}(S)_T C^z \in \mathbb{D}^{1,1}(\mathbb{R}^q)$ for all $z \in \mathbb{R}^q$ where C^z is given by (1.27) can be perfectly hedged*

$$X_t^{\hat{\pi},\hat{c}} = Y_t \;\; \mathbf{P} \otimes \boldsymbol{\lambda} \; a.e., \tag{1.28}$$

where $Y_t = (Y_t : t \in [0, T])$ denotes the value process of the contingent claim given by

$$Y_t = \mathbf{E}^{\hat{\mathbf{Q}}}[\frac{B_t}{B_T} H|\mathcal{G}_t], \tag{1.29}$$

and where $\tilde{\mathbf{Q}}$ corresponds to the insider's local martingale measure given by

$$\frac{d\tilde{\mathbf{Q}}}{d\mathbf{P}^z}_{|\mathcal{G}_t,z=G} = \frac{q(G)}{p(\omega,t,G)}\frac{d\mathbf{Q}}{d\mathbf{P}}_{|\mathcal{F}_t}.$$

Corresponding replicating strategies $(\hat{\pi},\hat{c})$ are given by

$$\hat{c}_t = e_t \tag{1.30}$$

for consumption and

$$(\hat{\pi}_t)^* = \mathcal{D}_t\mathbf{E}^{\mathbf{P}^z}[\frac{\mathcal{E}(S)_T}{\mathcal{E}(S)_t}C^z|\mathcal{F}_t]_{|\{z=G\}}\sigma_t^{-1} \tag{1.31}$$

for portfolio investments.

2. *For an investment horizon $T \in [\![T_G, 1]\!]$ the contingent claim has a positive implicit price if and only if insider information $\mathbb{G}$ is atomic.*

It follows from the expression for the hedging strategy (1.31) that the fraction of wealth that has to be invested in each risky asset is given by the contemporaneous elasticity of the value process (1.28) with respect to changes in the state of nature multiplied by the inverse of the elasticity of prices with respect to changes in the state of nature

$$\frac{\hat{\pi}_t}{X_t^{\hat{\pi},\hat{c}}} = ((\mathcal{D}_t[[\log P_t^j]]_{j=1,\dots,d})^*)^{-1}(\mathcal{D}_t \log Y_t)^*. \tag{1.32}$$

From lemma 23 and the proof of the theorem we see that the initial price an investor with an enlarged flow of information is willing to pay for the contingent claim given by $\mathbf{E}[\mathcal{E}(-\int_0^{\cdot}(\alpha_s^z)^* dW_s^G)_T\mathcal{E}(S)_T H|\mathcal{G}_0]$ corresponds to the price an investor who is "gambling" upon the event $\{G = z\}$ is willing to pay $\mathbf{E}[\mathcal{E}(S)C^z]$ evaluated at $z = G$. Furthermore as we show in the proof a non-insider faces a tracking error given by $\int_T^{T_G}\mathbf{E}[\mathcal{D}_u[\mathcal{E}(S)_T H|\mathcal{F}_u]d\omega(u)$ even if there are no constraints and markets are dynamically complete. This stems from the fact that the claim considered in the theorem is supposed to be $\mathcal{G}_T$- and not $\mathcal{F}_T$- measurable and therefore at maturity non-insiders do not know its actual pay-off with certainty.

1.4.2 Claims measurable with respect to Public Information

Claims such that even at maturity their pay-off is unknown for common available information exist whenever there are some non-traded goods in the economy. But the existence of a market for options with uncertain pay-offs even at maturity for some investors may be unrealistic. Independently, The results of the previous theorem have their own interest since they will be used for the solution of the consumption-investment problem for insiders in the next section.

If insiders' information is revealed through market prices in equilibrium, contingent claims written on prices will in fact be $\mathcal{G}_T$- and not $\mathcal{F}_T$- measurable. In this case theorem 10 provides the relevant pricing formulas and hedging strategies.

For contingent claims whose cash flow at maturity is known also for non-insiders we have the following result:

Theorem 11 *In the absence of an arbitrage opportunity an insider and a non-insiders have the same valuation for a claim $H \in \mathbf{L}^1(\Omega, \mathcal{F}_T, \mathbf{P})$. Its implicit price is given by*

$$Y_0 = \mathbf{E}[\mathcal{E}(S)_T H]. \tag{1.33}$$

It follows that if the contingent claim is $\mathcal{F}_T$- adapted its pay-off structure does not depend on the random element the flow of information is sharpened with. Its value is then the same for investors knowing the random vector G, and for investors who just have common available information to replicate the contingent claim. Obviously, replication costs of insiders are not higher than those of non-insiders since they can replicate the contingent claim by strategies that depend on the coarser common available flow of information.

The reason why prices of $\mathcal{F}_T$- measurable contingent claims are unaffected by insider information follows from the fact that absolutely continuous local martingale measure $\tilde{\mathbf{Q}}$ coincides on $\mathcal{F}_T$ for $T \in [\![0, T_G[\![$ with the equivalent local martingale measure $\mathbf{Q}$ of investors having just public information.

As we have seen in theorem 1 enlargements of filtrations do not affect the quadratic variation of the prices of risky assets. Consequently the invariance of the implicit price with respect to insider information is basically just a consequence of the well-known fact first discovered for the Black and Scholes formula that option prices do not depend on the drift coefficient of the risky assets. Since as we have seen initial enlargements of filtrations can be derived from a Girsanov transformation with respect to a conditional measure this result is not surprising since Girsanov transformations do not affect the quadratic variation which implies invariance of the option prices with respect to changes of measures and initial enlargements of filtrations.The invariance with respect to changes of filtrations and heterogeneity of equivalent beliefs will disappear if perfect replication of the claim is impossible or as we have seen if insider information is non-atomic such that no local martingale measure for the insider will exist. In this case the investor with insider information may be capable of constructing an investment strategy with smaller tracking error or with the same tracking error but higher a probability to replicate the pay-off structure of the claim than the non-insiders. To analyze such effects we need to relax the assumption that initial available information is given by a Wiener filtration.

Duffie and Huang (1986) have shown that it is necessary for a fully revealing rational expectation equilibrium to exist that the hedging costs of agents having ordered flows of information be the same. Otherwise better informed investors would have an arbitrage opportunity. Our result shows that if the difference of flows of information is atomic an insider and a non-insider agree on the implicit price of the contingent claim, independently of whether or not the insider takes into account that his/her information will be fully revealed through equilibrium prices. It follows that their conclusion, that the advantage of better information is not such that the costs of hedging are lower for the better informed, but such that the set of claims that can be replicated is enlarged, does not depend on the fact that the information about the insider signal is contained in equilibrium prices. Our analysis shows that it is already a consequence of the fact that the absolutely continuous local martingale measure of an insider does correspond to the risk neutral probability of a non-insider on public information.

On the other hand theorem 8 shows that the conclusions of Duffie and Huang are only valid if the differences in the information flows are atomic or never disappear. Then whenever an insider has non-atomic anticipative information and an investment horizon which includes the first moment in time his/her insider information will be known publicly, any contingent claim can be replicated with zero initial capital, and as a consequence asset prices in a fully revealing rational expectations equilibrium will necessarily be zero.

1.5 Insider Information, Portfolio and Consumption Policies

In this section we solve Merton's consumption-investment problem for an investor with information given by a Wiener filtration and additional information generated by a $\mathcal{F}_{T_G}$- measurable random vector G. Then we derive explicit expressions for portfolio policies, which show that the insider's demand for risky assets can be decomposed in a part which is dependent on the state price density, another part which depends on the endowment process and a last part which purely depends on his/her anticipative information. Finally, we analyze the information about insider signals contained in optimal strategies.

1.5.1 Merton's Problem for an Insider

Since we have shown in the previous section that smooth claims measurable with respect to enlarged information can be perfectly hedged we can generalize previous results of of Karatzas and Pikovsky (1996a), Elliott, Geman and Korkie (1997) and Amendinger, Imkeller and Schweizer (1998) and allow for more general preferences and for consumption before the final period.

The approach presented here differs from those previously considered by the choice of a conditional criterion function, that is we maximize expected utility of consumption conditional on initial information. Furthermore we assume that insiders do not care about final wealth. Given that initial Wiener filtrations are trivial this is equivalent to unconditional optimization if there is no side information. As it can easily be seen optimal policies for the conditional criterion must be optimal for the unconditional one. As a consequence the marginal value of wealth for insiders in our model will be state dependent.

In our model Merton's consumption investment problem with admissible strategies for an insider $(\pi, c) \in \mathcal{A}(\mathbf{P}, \mathbb{G}, e)$ can be formulated as follows.

$$\sup_{(\pi,c)\in\mathcal{A}(\mathbf{P},G,e)} \mathbf{E}[\int_0^T u(v, c_v)dv|\mathcal{G}_0]. \tag{1.34}$$

The following assumption is necessary for the existence of a solution.

Assumption 4

$$\mathcal{X}(y, z) < \infty \tag{1.35}$$

for all $z \in \mathbb{R}^q$ *where*

$$\mathcal{X}(y, z) := \mathbf{E}^{\mathbf{P}^z}[\int_0^T \frac{q(z)}{p(\cdot, t, z)}\mathcal{E}(S)_v I(t, y\mathcal{E}(S)_v \frac{q(z)}{p(\cdot, v, z)})dv] \tag{1.36}$$

Merton's consumption-investment problem without liquidity constraints in an economy with future endowments can be interpreted as a problem in an economy where initial liquid wealth is given by the implicit price of cumulative endowments at the initial date and no future endowments. If there are no arbitrage opportunities marketed wealth e^G at time zero for an insider is as follows

$$e^G = e_0 + \mathbf{E}^{\tilde{\mathbf{Q}}}[\int_0^T \frac{e_s}{B_s} ds|\mathcal{G}_0]. \tag{1.37}$$

Similarly, for conditional beliefs $\mathbf{P}^z$ we define the process $Y^z = (Y_t^z; t \in [0,T])$ by

$$Y_t^z := \mathbf{E}^{\mathbf{P}^z}[\int_t^T \frac{\mathcal{E}(S^z)_v}{\mathcal{E}(S^z)_t}(\hat{c}_v^z - e_v)dv|\mathcal{F}_t], \tag{1.38}$$

and where the conditional state price density $\mathcal{E}(S^z)$ is given by

$$\mathcal{E}(S^z)_t = \frac{q(z)}{p(\omega,t,z)}\mathcal{E}(S)_t. \tag{1.39}$$

The process Y^z corresponds to the implicit price of optimal net consumption $\hat{c}^z$ for an investor with beliefs given by the conditional Wiener measure.

The following theorem provides the solution of Merton's consumption-investment problem for an insider. It shows that for the existence of a viable market model without any restrictions on the investment horizon it is necessary that insider information be atomic. The anticipative information affects optimal strategies through the likelihood ratio between the conditional and unconditional density of the signal. It follows that an insider will never consume more or less in all states of nature than a non-insider.

Theorem 12 *Under the assumptions of theorem 1 we have the following:*

1. *An investor with investment horizon $T \in [\![T_G, 1]\!]$ who has non-atomic insider information attains ex ante infinite expected utility.*

2. *If the insider has only atomic anticipative information or the investment horizon terminates before the uncertainty about the signal is resolved* $T \in [\![0, T_G[\![$, *the solution to Merton's consumption-investment problem is if assumption 4 is satisfied as follows:*

The optimal consumption policy is given by:

$$\hat{c}_t = I(t, \hat{y}^G \frac{q(G)}{p(\omega, t, G)} \mathcal{E}(S)_t), \tag{1.40}$$

where $\hat{y}^G$ *corresponds to the marginal value of initial wealth and satisfies:*

$$\mathcal{X}(\hat{y}^G, G) = e^G. \tag{1.41}$$

The optimal portfolio policy is given by:

$$\hat{\pi}_t = \hat{\pi}_t^G \tag{1.42}$$

where for all $z \in \mathbb{R}^q$

$$(\hat{\pi}_t^z)^* = \mathcal{D}_t Y_t^z (\sigma_t)^{-1}. \tag{1.43}$$

The insider's optimal wealth $X_t^{\hat{\pi},\hat{c}}$ *satisfies*

$$X^{\hat{\pi},\hat{c}_t} = Y_t^G, \mathbf{P} \otimes \boldsymbol{\lambda} \ \ a.e.. \tag{1.44}$$

For logarithmic utility for final wealth only Amendinger, Imkeller and Schweizer (1998) show that the existence of a solution to Merton's problem depends on the relative entropy between the conditional and unconditional law of the signal. They show for these particular preferences that a solution exists whenever the relative entropy is finite. It follows from theorem 8 that if there is an arbitrage the conditional and unconditional laws of the signal at resolution time are mutually singular, and consequently the corresponding relative entropy is infinite.

Since as we have seen in theorem 8 arbitrage opportunities occur just immediately before a resolution time it follows that for investment horizons which end before there is no arbitrage.

If the investment horizon covers the resolution time it is necessary that all possible events revealed by the insider information are non-zero probability events. In this case additional information is not sufficiently fine to realize an arbitrage and therefore the demand for risky assets is not infinite.

The results of theorem 12 illustrate that even if there exists only an absolutely continuous but not equivalent local martingale measure, Merton's consumption-investment problem still has a solution. The static budget constraint associated with the dynamic problem with conditional beliefs and investment horizon T can be written

$$\mathbf{E}^{\mathbf{P}^z}[\mathbf{1}_{\{\frac{dP}{dP^z}|_{\mathcal{F}_T}\in]0,+\infty[\}}\frac{d\mathbf{P}}{d\mathbf{P}^z}|_{\mathcal{F}_T}\int_0^T \mathcal{E}(S)_t c_t dt] \leq e^z, \tag{1.45}$$

or equivalently

$$\mathbf{E}[\mathbf{1}_{\{\frac{dP}{dP^z}|_{\mathcal{F}_T}\in]0,+\infty[\}}\int_0^T \mathcal{E}(S)_t c_t dt] \leq e^z \tag{1.46}$$

since on $\{\frac{d\mathbf{P}}{d\mathbf{P}^z}|_{\mathcal{F}_T}\in]0,+\infty[\}$ we have that $\mathbf{P}^z \sim \mathbf{P}$. This proves that though $\mathbf{P}(\{\frac{d\mathbf{P}}{d\mathbf{P}^z}|_{\mathcal{F}_T}=+\infty\})>0$ and $\mathbf{P}^z(\{\frac{d\mathbf{P}}{d\mathbf{P}^z}|_{\mathcal{F}_T}=+\infty\})=0$ for $T\in[\![T_G,1]\!]$ binary options $\mathbf{1}_{\{\frac{dP}{dP^z}|_{\mathcal{F}_T}=+\infty\}}$ will not add current consumption without violating the budget constraint [14].

[14] This seems to contradict results in Dybvig and Willard (1996). They claim that constraints, which prevent "empty promises" in states which are not believed to occur given the public beliefs, are necessary for the existence of a solution to the static consumption-investment problem associated with the dynamic problem. They assume that there are no future endowments. The difference with our results arises since they consider the budget constraint under the common beliefs $\mathbf{P}$, given by $\mathbf{E}[\int_0^T \mathcal{E}(S)_t c_t dt]=e_0$, which would correspond to the static budget constraint associated with the dynamic consumption-investment problem only if $\mathbf{P}\sim\mathbf{P}^z$ on $\mathcal{F}_T$ for all $T\in[0,1]$. Now suppose there exists a $\mathcal{F}_t$- adapted consumption policy $\tilde{c}_t$ which satisfies the

1.5.2 Optimal Portfolio Policies for Insiders

Optimal investment policies of insiders (1.43) in theorem 12 are given as Malliavin derivative of the respective implicit price of net-consumption of an investor who "gambles" upon the event $\{G = z\}$ evaluated at the true realization of the signal. It follows that if we can find an explicit expression for optimal wealth for an investor with beliefs given by the conditional Wiener measure $\mathbf{P}^z$, we are also capable to get explicit expressions for the optimal investment policies of an insider since these are given as instantaneous Malliavin derivatives of "gambler's" optimal wealth evaluated at the true signal. This result is similar to the way optimal investment policies without insider information are obtained in Markov markets as described by Cowell, Elliott and Kopp (1991) and Benoussan and Elliott (1995).

Under a Markovian assumption, optimal wealth is given as a function of state variables and the corresponding optimal portfolio policy is derived with Itô's rule from the derivatives of this function. Our setup is non-Markovian and we obtain optimal portfolio policies from Malliavin derivatives of the optimal wealth with respect to states of nature. The Wiener processes play in a non-Markovian market

budget constraint for public beliefs and is such that

$$\mathbf{E}^{\mathbf{P}^z}[\int_0^T u(t, \tilde{c}_t)dt] > \mathbf{E}^{\mathbf{P}^z}[\int_0^T u(t, \hat{c}_t^z)dt].$$

It follows from Lebesgue's decomposition

$$\mathbf{E}[\int_0^{T_G} \mathcal{E}(S)_t \tilde{c}_t dt] = \mathbf{E}^{\mathbf{P}^z}[\frac{d\mathbf{P}}{d\mathbf{P}^z}_{|\mathcal{F}_{T_G}} \int_0^{T_G} \mathcal{E}(S)_t \tilde{c}_t dt] + \mathbf{E}[\mathbf{1}_{\{\frac{dP}{dP^z}|_{\mathcal{F}_{T_G}} = +\infty\}} \int_0^{T_G} \mathcal{E}(S)_t \tilde{c}_t dt]$$

that the consumption policies $\tilde{c}$ will also satisfy the budget constraint for the conditional beliefs. The existence of such a consumption policy contradicts therefore the optimality of $\hat{c}_t^z$. As a consequence the value of the static optimization problem with feasible consumption policies for the public beliefs must be bounded from above by the value of the static problem given conditional beliefs. This illustrates that in our model non-empty promises constraints will not be binding.

the role of prices in the determination of optimal portfolio policies and the "delta hedging term" is given as an instantaneous Malliavin derivative of the claim's implicit price with respect to the Wiener processes. Malliavin derivatives measure the "contemporaneous sensitivity" of wealth with respect to changes in states of natures.

To get explicit expressions for the conditional expectation in the expression for the optimal wealth process in a non-Markovian setup may be difficult. In this case, to get more explicit expressions for optimal portfolio policies, we can interchange the conditional expectation and derivative operator to get representations similar to those obtained by Karatzas and Ocone (1991) for Brownian filtrations. This allows us to analyze how investment strategies depend on anticipative information.

The following expression shows that portfolio policies depend on the "sensitivity" of optimal consumption policies with respect to changes in states of nature (i.e. Malliavin derivatives).

$$\begin{aligned}(\hat{\pi}_t^G)^* \sigma_t = \mathbf{E}[\int_t^T \frac{\partial_2 u(v, \hat{c}_v^G)}{\partial_2 u(t, \hat{c}_t^G)} (\hat{c}_v^G - e_v) dv | \mathcal{G}_t] A(t, \hat{c}_t^G) (\mathcal{D}_t \hat{c}_t^z)_{|z=G} \\ + \mathbf{E}[\int_t^T \frac{\partial_2 u(v, \hat{c}_v^G)}{\partial_2 u(t, \hat{c}_t^G)} [1 - (\hat{c}_v^G - e_v) A(v, \hat{c}_v^G)] (\mathcal{D}_t \hat{c}_v)_{|z=G} dv | \mathcal{G}_t] \\ - \mathbf{E}[\int_t^T \frac{\partial_2 u(v, \hat{c}_v^G)}{\partial_2 u(t, \hat{c}_t^G)} \mathcal{D}_t e_v) dv | \mathcal{G}_t] \quad (1.47)\end{aligned}$$

The instantaneous "sensitivity" $(\mathcal{D}_t \hat{c}_t^z)_{|z=G}$ is related to an expression similar to the CCAPM equation, then:

$$\mathcal{D}_t \hat{c}_t^z = \partial_2 I(t, \hat{y}^z \mathcal{E}(S^z)_t) \hat{y}^z \mathcal{E}(S_t^z) \mathcal{D}_t S_t^z, \quad (1.48)$$

where we have used that $\mathcal{D}_t [S^z, S^z]_t = 0$. Since

$$\partial_2 I(t, \hat{y}^z \mathcal{E}(S^z)_t) \hat{y}^z \mathcal{E}(S_t^z) = -\frac{1}{A(t, \hat{c}_t^z)}, \quad (1.49)$$

where $A(t,c)$ denotes absolute risk aversion and

$$\mathcal{D}_t S_t^z = -(\alpha_t^z + \theta_t^z)^*, \tag{1.50}$$

we see that the conditional Sharpe ratio $\theta + \alpha^G$ for an insider is proportional to the "sensitivity" of optimal consumption of an investor who conditions on the event $\{G = z\}$ with respect to changes in states of nature, evaluated at the true signal. The proportionality factor is determined by the insider's absolute risk aversion.

$$[A(t, \hat{c}_t^z)(\mathcal{D}_t \hat{c}_t^z)^*]_{|z=G} = \alpha_t^G + \theta_t. \tag{1.51}$$

If we calculate all the Malliavin derivatives of consumption and endowment rate processes for insiders, we can break down the individual demand for risky assets into three components: one depending on state price density, another depending on the individual endowment process and a component depending on the anticipative information.

Proposition 13 *Under the assumptions of theorem 12 we have that the insider's demand for risky assets $\hat{\pi} = \hat{\pi}^G$ can be written as follows*

$$\hat{\pi}_t^G = \hat{\pi}_t^{G,S} + \hat{\pi}_t^{G,E} + \hat{\pi}_t^{G,I}, \tag{1.52}$$

as long as $\mathcal{E}(S)_T X_T^{\hat{\pi},\hat{c}} \in \mathbb{D}^{1,1}(\mathbb{R}^d)$. The demand $\hat{\pi}^{G,S}$ denotes the demand for risky assets that depends on the sensitivity of the state price density with respect to changes in the state of nature, $\hat{\pi}^{G,E}$ corresponds to the demand that depends on the sensitivity of the endowment process with respect to states of nature and $\pi^{G,I}$ describes the demand which depends on the insider signal G.

In what follows we give explicit expressions for the different parts of the demand for risky assets. As we mentioned before optimal policies of an insider can be derived as those of an investor who conditions on $\{G = z\}$ for some $z \in \mathbb{R}^d$. Such an investor has a financial market model given by $(\Omega, \mathcal{F}_1, \mathbb{F}, \mathbf{P}^z, \mathcal{C}, \succeq)$. We see that his/her beliefs correspond to the conditional Wiener measure derived in Appendix 1.7. The consumption-investment problem of such an investor can be derived for atomic insider signals on the stochastic basis for non-insiders if we introduce the state dependent utility function

$$v(t, c, \omega, z) := \frac{p(\omega, t, z)}{q(z)} u(t, c). \tag{1.53}$$

The "discount factor" $\frac{p(\omega,t,z)}{q(z)}$ in this representation corresponds to the density process $\frac{d\mathbf{P}^z}{d\mathbf{P}}_{|\mathcal{F}_t}$. With this utility function it is possible to represent agents with heterogeneous beliefs on the same stochastic basis.

Corollary 14 *The demand $\pi^{G,S}$ in proposition 13 that depends on the state price density has itself a "myopic" and a "dynamic" component*

$$\hat{\pi}_t^{G,S} = \hat{\pi}_t^{G,S,m} + \hat{\pi}_t^{G,S,d}. \tag{1.54}$$

The "myopic" component given by

$$\hat{\pi}_t^{G,S,m} := X_t^{\hat{\pi}^G, \hat{c}^G} (\sigma_t \sigma_t^*)^{-1} (b_t - 1_d r_t), \tag{1.55}$$

and the "dynamic" component $\hat{\pi}^{G,S,d}$ is such that on $\{z = G\}$

$$\hat{\pi}_t^{z,S,d} := (\sigma_t^*)^{-1} \mathbf{E}[\int_t^T \frac{\partial_2 v(u, \hat{c}_u^z, \cdot, z)}{\partial_2 v(t, \hat{c}_t^z, \cdot, z)} [e_u - \hat{c}_u^z + \frac{1}{A(u, \hat{c}_u^z)}] (\mathcal{D}_t \log \mathcal{E}(S)_u)^* du | \mathcal{F}_t], \tag{1.56}$$

where the logarithmic Malliavin derivative of the state price density is given by

$$(\mathcal{D}_t \log \mathcal{E}(S)_u)^* = -[\theta_t + \int_t^u (\mathcal{D}_t r_s)^* ds + \int_t^u (\mathcal{D}_t \theta_s)^* (\theta_s ds + d\omega(s))]. \tag{1.57}$$

The myopic component describes the demand of an insider with logarithmic utility function. This part depends on the deflator process S. We see that whenever the state price density is Gaussian the hedging demand is zero.

A similar decomposition exists for the demand that depends on the endowment rate process. The demand depends on the sensitivity of endowment with respect to states of nature. Since the endowment rate process is given exogenously and therefore given as a stochastic differential equation we obtain more explicit expressions for Malliavin derivatives.

Corollary 15 *The demand $\pi^{G,E}$ in proposition 13 that depends on the sensitivity of the endowment rate process with respect to changes in the state of nature is such that on $\{G = z\}$*

$$\hat{\pi}_t^{z,E} := -(\sigma_t^*)^{-1}\mathbf{E}[\int_t^T \frac{\partial_2 v(u, \hat{c}_u^z, \cdot, z)}{\partial_2 v(t, \hat{c}_t^z, \cdot, z)}(\mathcal{D}_t e_u)^* du|\mathcal{F}_t], \tag{1.58}$$

where the Malliavin derivative of the endowment rate process satisfies the diffusion for the endowment rate process with linearized coefficients.

$$(\mathcal{D}_t e_u)^* = \gamma^e(t, e_t) + \int_t^u (\mathcal{D}_t e_s)^*[(\partial_2\mu^e(s, e_s))ds + (\partial_2\gamma^e(s, e_s))^* d\omega(s)]. \tag{1.59}$$

Finally we present the part of the demand in the decomposition that determines how the optimal demand depends on the anticipative information. The demand which depends on the market price of risk $\pi^{G,S}$ depends on the change of measure $\frac{d\mathbf{Q}}{d\mathbf{P}}$ which determines the martingale measure of a non-insider. Similarly we see below that the demand which depends exclusively on insider information depends on the density process of conditional and unconditional law of the signal.

Corollary 16 *The demand $\pi^{G,I}$ in proposition 13 that depends on the insider's signal has a "myopic" and a "dynamic" component*

$$\hat{\pi}_t^{G,I} = \hat{\pi}_t^{G,I,m} + \hat{\pi}_t^{G,I,d}, \tag{1.60}$$

where the "myopic" component of the demand is given by

$$\hat{\pi}_t^{G,I,m} = X_t^{\hat{\pi}^G,\hat{c}^G}(\sigma_t^*)^{-1}\alpha_t^G, \tag{1.61}$$

whereas the "dynamic" component of the demand $\hat{\pi}^{G,I,d}$ *is such*

$$\hat{\pi}_t^{z,I,d} := (\sigma_t^*)^{-1}\mathbf{E}[\int_0^T \frac{\partial_2 v(u,\hat{c}_u^z,\cdot,z)}{\partial_2 v(t,\hat{c}_t^z,\cdot,z)}[\hat{c}_u^z - \dot{e}_u - \frac{1}{A(u,\hat{c}_u)}]\frac{\mathcal{D}_t p(\omega,u,z)}{p(\omega,u,z)}du|\mathcal{F}_t], \tag{1.62}$$

We see that only the myopic demand for risky assets [15] of the optimal portfolio policies can be unambiguously signed. The myopic demand for risky assets is increasing in the conditional spread between risky and risk free assets with respect to the conditional Wiener measure. The insider increases his/her demand for risky assets whenever risky assets and the conditional density of the signal are locally positively correlated.

Since $\alpha_t^z = 0$ for all $t \in [\![T_G, 1]\!]$ we see that after the resolution the insider's strategies only differ from $\mathcal{F}_t$- adapted optimal policies trough different optimal consumption processes.

From the expression for the optimal portfolio policies we can see how different assumptions can considerably simplify terms.

1. We obtain the demand for risky assets with non-anticipative information by putting $\alpha_t^z = 0$, which implies $\hat{\pi}_t^{\cdot,I,\cdot} = 0$ and $v = u$.

2. For deterministic coefficients ("Gaussian model") the Malliavin derivatives $\mathcal{D}_t\theta_s = 0$ and $\mathcal{D}_t r_s = 0$ and therefore

$$\hat{\pi}_t^{z,S} = \mathbf{E}[\int_t^T \frac{\partial_2 v(u,\hat{c}_u^z,\cdot,z)}{\partial_2 v(t,\hat{c}_t^z,\cdot,z)}\frac{1}{A(u,\hat{c}_u^z)}du|\mathcal{F}_t](\sigma_t\sigma_t^*)^{-1}(b_t - 1_d r_t), \tag{1.63}$$

and the corresponding part of the demand for insiders in Gaussian models is found by $\hat{\pi}_t^{G,S} = (\hat{\pi}_t^{z,S})_{|z=G}$. Since the Malliavin derivatives of Gaussian

[15] By "myopic" we mean the demand of an investor with logarithmic Bernoulli indicator.

random variables are deterministic, Gaussianity is a strong simplifying assumption.

3. Utility functions of the HARA type simplify expressions since these satisfy by definition $A(t,c)^{-1} = (1-\nu(t))c + \eta(t)$ and as a consequence:

$$\hat{\pi}_t^{z,S,d} = (\sigma_t^*)^{-1}\mathbf{E}[\int_t^T \frac{\partial_2 v(u,\hat{c}_u^z,\cdot,z)}{\partial_2 v(t,\hat{c}_t^z,\cdot,z)} \times [e_u + \nu(u)\hat{c}_u^z + \eta(u)](\mathcal{D}_t \log \mathcal{E}(S)_u)^* du|\mathcal{F}_t],$$

whereas in this case,

$$\hat{\pi}_t^{z,I,d} = (\sigma_t^*)^{-1}\mathbf{E}[\int_t^T \frac{\partial_2 v(u,\hat{c}_u^z,\cdot,z)}{\partial_2 v(t,\hat{c}_t^z,\cdot,z)}[e_u - (1-\nu(u))\hat{c}_u^z + \eta(u)] \times (\mathcal{D}_t \log \mathcal{E}(-\int_0^{\cdot} (\alpha_s^z)^* dW_s^z)_u)^* du|\mathcal{F}_t].$$

Corresponding demands for insiders are obtained if these expressions are evaluated at $z = G$. The most simple case in this class is logarithmic utility $u(t,c) = h(t)\log c$ since then $(A(t,c))^{-1} = c$ and in the above expressions $\nu = 1$ and $\eta = 0$.

4. If endowments are deterministic $e_t = e(t)$ for all $t \in [0,1]$ we have that $\mathcal{D}_s e_t = 0$ and as a consequence $\hat{\pi}_t^{\cdot,E} = 0$.

5. The simplest expression is obtained if we assume logarithmic utility and deterministic endowments then in this case we have just myopic demand for risky assets $\hat{\pi}_t^{\cdot} = \hat{\pi}_t^{\cdot,S,m} + \hat{\pi}_t^{\cdot,I,m}$. This case with the additional assumption of preferences for final wealth was the only one considered in Karatzas and Pikovsky (1996a), Elliott, Geman and Korkie (1997) as well as Amendinger, Imkeller and Schweizer (1998) .

1.5.3 The Value of Insider Information

To rationalize ex-ante the use of additional side information is easy. A finer flow of information reduces at each moment in time uncertainty about the true state of nature. Since investors are risk averse they will like this. That they indeed do can be seen from a comparison of the value functions for a normal investor with common beliefs and information

$$V(\mathbf{P}, \mathbb{F}, e) := \sup_{(\pi,c)\in\mathcal{A}(\mathbf{P},F,e)} \mathbf{E}[\int_0^T u(t, c_t)dt], \tag{1.64}$$

with that of an insider

$$V(\mathbf{P}, \mathbb{G}, e) := \sup_{(\pi,c)\in\mathcal{A}(\mathbf{P},G,e)} \mathbf{E}[\int_0^T u(t, c_t)dt]. \tag{1.65}$$

Since an insider can always choose his optimal strategies to be just $\mathcal{F}_t$- adapted we must by optimality of his strategies have that

$$V(\mathbf{P}, \mathbb{G}, e) - V(\mathbf{P}, \mathbb{F}, e) \geq 0. \tag{1.66}$$

It follows that the left hand side of this inequality can be interpreted as the individual value of better information.

In section 1.3 we have seen that insiders do not change the representation of market data on the stochastic basis $(\Omega, \mathcal{F}_1, \mathbb{G}, \mathbf{P})$ which is relevant for their decisions if and only if their anticipative information is independent of public information $G \perp \mathcal{F}_t$ for all $t \in [\![0, T_G[\![$. An equivalent result holds for the individual value of side information

Proposition 17 *Information generated by a random element G that satisfies assumption 1, 2 and 3 does not have any individual value:*

$$V(\mathbf{P}, \mathbb{G}, e) = V(\mathbf{P}, \mathbb{F}, e), \tag{1.67}$$

if it is independent of the public flow of information

$$G \perp \mathcal{F}_t, \quad \textit{for all } \; t \in [\![0, T_G[\![. \tag{1.68}$$

Then it is never possible to ex-ante Pareto improve an allocation by allowing the optimal strategies to depend upon information that cannot be learned from the common flow of information. For this to happen we would need for at least one investor a strict inequality between the value functions with and without side information.

1.5.4 The Informational Content of Insider Portfolio and Consumption Policies

For many equilibrium results it is crucial to determine whether optimal insider strategies reveal all private information or not. To show that optimal strategies are fully revealing it is necessary that the filtration generated by the optimal strategies correspond to the insider's flow of information. The following proposition shows that optimal portfolio and consumption policies are never fully revealing if insider information is given by initial enlargements of Wiener filtration.

Proposition 18 *If the insider information in non-redundant, optimal insider strategies (1.40) and (1.43) in theorem 12 are never fully revealing, since*

$$\mathcal{G}_t \not\subset \mathcal{F}_t^{\hat{c}^G} \bigvee \mathcal{F}_t^{\hat{\pi}^G} \tag{1.69}$$

for some $t \in [0, T]$.

This result shows that if the insider can only reveal his superior information through his state contingent investment and consumption demand, rational expectation equilibria in our model will never be fully revealing. A Walrasian auctioneer or market maker who just obtains the contingent investment demand in the form of any kind of market orders will never be able to fully infer the information about the state of nature contained in the investor's anticipative signals. For a fully revealing equilibrium to exist it is necessary that insiders communicate their information through other channels.

1.6 Two Examples of Insider Information

In this section we consider first an investor who already knows at the beginning of the investment horizon the time at which a stock stops to pay a dividend. Then we will derive optimal portfolio strategies for an investor who knows already at the beginning whether or not a dividend will be paid after a certain time. This kind of insider information has previously not been considered in the literature.

For simplicity we consider in both examples the following dividend process

$$dD_t = D_t[-2\gamma D_t^{-\frac{1}{2\gamma}} dW_t - 2\gamma(1-2\gamma) D_t^{-\frac{1}{\gamma}} dt] \text{ with } D_0 = a^{2\gamma}, \tag{1.70}$$

some $\gamma \geq 1$ and $a > 0$.

1.6.1 Non-Atomic Insider Information

Since the solution of this stochastic differential equation is $D_t = (a - W_t)^{2\gamma}$ an investor who knows at the beginning whether or not the stock does pay a dividend up to a certain moment in time has anticipative information generated by the signal $G = \inf\{t : D_t = 0\}$. Since $\{D_t = 0\} = \{W_t = a\}$ the signal G corresponds to the passage time $T_a = \inf\{t : W_t = a\}$ of the Brownian motion W at a. It is well known (see Karatzas and Shreve (1987) proposition 8.2 p.96) that the density of the passage time of a Brownian motion starting at zero T_a has the following density

$$\mathbf{P}^0(T_a \in dz) = \frac{a}{\sqrt{2\pi z^3}} \exp\{-\frac{a^2}{2z}\} dz, \tag{1.71}$$

and consequently it follows from the Markovian property of Brownian motion and the expression for the solution D_t that in our context $p(\omega, t, z) =: \tilde{p}(D_t(\omega), t, z)$ with

$$\tilde{p}(d,t,z) = \frac{d^{\frac{1}{2\gamma}}}{\sqrt{2\pi(z-t)^3}} \exp\{-\frac{d^{\frac{1}{\gamma}}}{2(z-t)}\}\mathbf{1}_{z>t}, \tag{1.72}$$

and therefore the logarithmic Malliavin derivative is given as follows $\alpha_t^z(\omega) =: \tilde{\alpha}(D_t, t, z)$ where

$$\tilde{\alpha}(d,t,z) = (-\frac{1}{d^{\frac{1}{2\gamma}}} + \frac{d^{\frac{1}{2\gamma}}}{z-t})\mathbf{1}_{z>t}, \tag{1.73}$$

and consequently

$$\alpha_t^G = (\frac{1}{D_t^{\frac{1}{2\gamma}}} + \frac{D_t^{\frac{1}{2\gamma}}}{G-t})\mathbf{1}_{G>t}. \tag{1.74}$$

If we use this expression in theorem 1 we get the model relevant for an insider knowing exactly the time of the last dividend payment. By left continuity of α_t^G we $\mathbf{P}$-a.s. have that $(\alpha_{G-}^G)^2 = +\infty$ and consequently that the wealth process corresponding to the myopic strategy $\frac{\tilde{\pi}_t^G}{X^{\tilde{\pi}^G}} = \frac{\alpha_t^G}{\sigma_t}\mathbf{1}_{t<G}$; the corresponding wealth

$$X_t^{\tilde{\pi}^G,e} = \frac{\exp\{-\int_0^t \alpha_s^G \theta_s ds\}}{\mathcal{E}(-\int_0^{\cdot} \alpha_s^G dW_s^G)_t}, \tag{1.75}$$

converges $\mathbf{P}$-a.s. to $X_G^{\tilde{\pi}^G,e} = +\infty$. Clearly such an investor has an arbitrage opportunity given by $\tilde{\pi}$. The arbitrage is due to the fact that the information which is generated by the signal consists of events $\{T_a = z\}$ for some $z \in \mathbb{R}^d$ and is so precise that it contains events which are not believed to be a possible occurrence given initial public information $\mathbf{E}[\mathbf{1}_{\{T_a=z\}}] = 0$.

1.6.2 Atomic Insider Information

To illustrate the difference between atomic and non-atomic insider information we consider an insider who knows already at the beginning whether or not the last dividend payment will be after a certain time $T^\star$. The signal of such an insider is therefore $G(\omega) = \mathbf{1}_{]T^\star,+\infty[}(T_a(\omega))$. The information revealed by this signal is equivalent to the information whether or not the minimal dividend rate during the period $[0,T^\star]$ is positive. The important difference relative to the previous example is that the event $\{G = z\}$ is of positive probability when $z \in \{0,1\}$. The interpretation of this is that the insider signal contains information which was considered a possible occurrence given initial public information and in this sense contains less information about the true state of nature. The conditional density in this case is obtained from

$$p(\omega,t,z) = \int_{T^\star}^{+\infty} \tilde{p}(D_t(\omega),t,x)dx\mathbf{1}_{t<T^\star}, \tag{1.76}$$

which gives $p(\omega,t,z) =: \hat{p}(D_t(\omega),t,z)$ where for $t < T^\star$

$$\hat{p}(d,t,z) = ((2\Phi(\frac{d^{\frac{1}{2\gamma}}}{\sqrt{T^\star - t}}) - 1)\mathbf{1}_{\{1\}}(z) + 2(1-\Phi(\frac{d^{\frac{1}{2\gamma}}}{\sqrt{T^\star - t}})\mathbf{1}_{\{0\}}(z)), \tag{1.77}$$

where Φ denotes the cumulative Gaussian distribution with corresponding kernel ϕ. Consequently taking Malliavin derivatives we obtain in this case $\alpha_t^G(\omega) =: \hat{\alpha}(D_t(\omega),t,G)$ where

$$\hat{\alpha}(D_t,t,G) = \frac{\frac{1}{\sqrt{T^\star - t}}\phi(\frac{D_t^{\frac{1}{2\gamma}}}{\sqrt{T^\star - t}})}{\Phi(0)^G - \Phi(\frac{D_t^{\frac{1}{2\gamma}}}{\sqrt{T^\star - t}})}\mathbf{1}_{t<T^\star}. \tag{1.78}$$

If we put these expressions into those of theorem 1 we obtain again the model relevant for an insider who knows whether or not the dividend of a stock is always positive up to a given date. We see that for states such that the dividend is strictly positive up to $T^\star$ the conditional return of the stock is reduced. For states such that dividends fall to zero for the first time before $T^\star$ the conditional appreciation rate of stocks increases.

The events revealed by such a signal are countable and it follows from theorem 8 that there is no arbitrage for insiders whenever there is no arbitrage for non-insiders in this case. To focus on the portfolio strategies which do just depend on the information we assume for simplicity in what follows that the market price of risk $\theta_t = 0$, the spot rate $r_t = 0$ and the endowment rate process $e_t = 0$ all the time. In such a model non-insiders have in contrast to insiders no demand for risky assets. Therefore, the only demand for risky assets stems from the insider's anticipative information. Furthermore we assume that the utility function is from the CRRA family $u(t,c) = e^{-\rho t}\log c$ if the relative risk aversion is one $R = 1$ and $u(t,c) = e^{-\rho t}\frac{c^{1-R}}{1-R}$ else. Using the results of corollary 16 we obtain the myopic part of the demand for the stock as follows

$$\hat{\pi}_t^{G,I,m} = X_t^{\hat{\pi}^G,\hat{c}^G}\frac{\hat{\alpha}(\omega(t),t,G)}{\sigma_t}\mathbf{1}_{t<T^\star}, \tag{1.79}$$

The myopic part is such that on states for which dividends remain positive up to $T^\star$ he wants to short the stock. In contrast he wants to hold a long position in the stock if the dividend falls to zero before $T^\star$. Such a behavior is only characteristic for an investor with logarithmic utility. For other non-myopic preferences we have for $t < T^\star$ the following additional demand for risky assets

$$\hat{\pi}_t^{G,I,d} = \frac{-(1-1/R)}{\sigma_t}\int_t^{T^\star} dv\int_0^{+\infty}\hat{c}(y,v,G)\hat{\alpha}(y,v,G)d\Phi\left(\frac{D_t^{\frac{1}{2\gamma}} - y^{\frac{1}{2\gamma}}}{\sqrt{v-t}}\right). \tag{1.80}$$

where the $\hat{c}(D_t,t,G)$ corresponds to the optimal consumption policy of the insider given by $\hat{c}(d,t,z) = I(t,\hat{y}^z\frac{q(z)}{\hat{p}(d,t,z)})$. An insider who is less risk averse than an insider with logarithmic preferences will have higher short or long positions than a myopic insider. In contrast an insider who is more risk averse than a myopic insider want to hedge their short and long positions. Their net demand in the two cases cannot be signed. For more general Markovian setups, where explicit expressions cannot be found, the Monte Carlo techniques presented in Detemple, Garcia and Rindisbacher (1998) can be extended to study the different components of the portfolio strategies in proposition 13. Our example illustrates that the effects of anticipative information on optimal portfolio policies can depend on the assumptions about preferences. The decomposition of portfolio policies in this paper helps to clarify this dependence on preferences. Similarly, they show whether or not a conclusion is robust with respect to the probabilistic structure of the insider signal.

1.7 Appendix A:The "Girsanov Approach" to Initial Enlargements of Wiener filtrations

The aim of this appendix is to derive in detail all necessary results to get the semi-martingale representation of the wealth process on the stochastic basis $(\Omega, \mathcal{F}_1, \mathbb{G}, \mathbf{P})$ relevant for the choice of consumption and portfolio policies of an insider. In contrast to the filtering problem where Stricker's theorem guarantees that any $\mathbb{G}$ semi-martingale is also a $\mathbb{F}$- semi-martingale, stability of semi-martingales with respect to enlargements of filtration does not necessarily hold (see Jacod (1979) for a discussion). The presentation is based on Jacod (1980) , Foellmer and Imkeller (1993) and Imkeller (1996), but allows for enlargements with respect to random variables which are measurable with respect to Brownian filtration before the terminal date. We first derive a conditional Wiener measure and find semi-martingale decompositions for this measure. Corresponding compensators depend on a parameter which represents the realization of the signal. Using results in Stricker and Yor (1978) (proposition 2 and théorème 1) we can always pick a version which depends measurably on this parameter. We will always consider this version of such processes in what follows. The general idea behind the "Girsanov approach" to the enlargements of filtrations is due to Song (1987) . He also shows how the same idea can be exploited for progressive enlargements of filtrations.

A sufficient condition for the existence of a semi-martingale representation for initial enlargements ("hypothèse H") is Jacod's criterion ("condition A") (see Jacod (1980) page 15), which in our model can be stated as follows

Assumption 5 *("Condition A") There exists a common measure $\boldsymbol{\nu}$ on $\mathcal{B}_{R^q}$ such that $\mathbf{P}_t^\omega \ll \boldsymbol{\nu}$ for all $t \in [\![0, T_G[\![$ where $\mathbf{P}_t^\omega$ corresponds to the conditional law of G given the initial filtration $\mathcal{F}_t$.*

Remarks

1. If "condition A" is satisfied we can without loss of generality choose $\boldsymbol{\nu} = \mathbf{P}_G$.

2. Since for $t \in [\![T_G, 1]\!]$ the conditional law of the signal corresponds to $\mathbf{1}_{\{z\}}(G(\omega))$ "condition A" cannot be satisfied after resolution time unless G is constant. Furthermore it follows that if $\sigma(G)$ does not contain any atom then not only $\mathbf{P}_G \not\ll \mathbf{P}_t^\omega$ but even $\mathbf{P}_G \perp \mathbf{P}_t^\omega$ for $t \in [\![T_G, 1]\!]$. This follows since the conditional law takes all its positive mass on $\{G = z\}$, a null set of the unconditional law.

3. Imkeller (1996) has recently given sufficient conditions for which random variables $G \in \mathbb{D}^{1,2}(\mathbb{R}^q)$ on Wiener space have conditional laws that are absolutely continuous with respect to the Lebesgue measure $\boldsymbol{\lambda}$.

We now derive the compensator of a $\mathbb{F}$- martingale on $(\Omega, \mathcal{F}_1, \mathbb{G}, \mathbf{P})$. As it was sketched by Jacod (1980) and shown by Foellmer and Imkeller (1993) this can be done by means of a Girsanov transformation with a conditional (Wiener) measure. Since this approach provides the key to derive optimal strategies for an insider we present it in detail.

To understand the link between the conditional and unconditional law with the Wiener measure and conditional Wiener measure we introduce the joint law of insider signal G and states of nature W

$$B \times E \mapsto \mathbf{R}(B, E); \quad B \times E \in \mathcal{F}_1 \otimes \sigma(G), \tag{1.81}$$

where

$$\mathbf{R}(B, E) := \mathbf{E}[\mathbf{1}_{G^{-1}(B)} \mathbf{1}_E]. \tag{1.82}$$

Lemma 19 *1. For all $t \in [0,1]$ the measure*

$$\mathbf{P}_t^\omega(dz) := \frac{d\mathbf{R}}{d\mathbf{P}}_{|\mathcal{F}_t}(\omega, dz) \tag{1.83}$$

on $\sigma(G)$ corresponds to the conditional law of G given $\mathcal{F}_t$

2. *For all $t \in [0,1]$ the probability measure*

$$\mathbf{P}^z(d\omega) := \frac{d\mathbf{R}}{d\mathbf{P}_G}(d\omega, z) \tag{1.84}$$

corresponds to the Wiener measure restricted to $\{G = z\}$.

3. *For all $t \in [0,1]$ we have*

$$\frac{d\mathbf{P}^z}{d\mathbf{P}}_{|\mathcal{F}_t}(\omega) = \frac{d\mathbf{P}_t^\omega}{d\mathbf{P}_G}(z) \tag{1.85}$$

$\mathbf{P} \otimes \mathbf{P}_G$- *a.e..*

Proof

Since $\mathbf{P}(E) = \mathbf{R}(\mathbb{R}^q, E)$, respectively $\mathbf{P}_G(B) = \mathbf{R}(B, \Omega)$ it follows that $\mathbf{R}(\cdot, E) \ll \mathbf{P}_G$ for all $E \in \mathcal{F}_t$, respectively $\mathbf{R}(B, \cdot) \ll \mathbf{P}$ for all $B \in \sigma(G)$, we have that $\mathbf{P}_t^\omega$ respectively $\mathbf{P}^z$ are both absolutely continuous with respect to $\mathbf{P}$ and therefore by the Radon-Nikodym theorem that $\mathbf{P}_t^\omega$ and $\mathbf{P}^z$ exist and are given as Radon-Nikodym derivatives of $\mathbf{R}$ with respect to $\mathbf{P}$.

Furthermore since

$$\mathbf{P}^z(d\omega)\mathbf{P}_G(dz) = d\mathbf{R}(dz, d\omega) = \mathbf{P}_t^\omega(dz)\mathbf{P}(d\omega), \tag{1.86}$$

it follows that (1.85) must hold $\mathbf{P} \otimes \mathbf{P}_G$- a.e..

Q.E.D.

Remarks

1. The conditional Wiener measure $\mathbf{P}^z$ can be interpreted as beliefs of an investor “gambling” upon the event $\{G = z\}$.

2. The equality in (1.85) allows to represent processes relevant for decisions conditional on the event $\{G = z\}$ in terms of properties of the signal's conditional law only.

3. If "condition A" holds we have for $t \in [\![0, T_G[\![$ that

$$\mathbf{P}_t^\omega(dz) = p(\omega, t, z)\boldsymbol{\nu}(dz), \tag{1.87}$$

and consequently that

$$\frac{d\mathbf{P}^z}{d\mathbf{P}}_{|\mathcal{F}_t}(\omega) = \frac{p(\omega, t, z)}{q(z)}, \tag{1.88}$$

where $q(z) := p(\omega, 0, z)$ corresponds to the density of $\mathbf{P}_G$ with respect to $\boldsymbol{\nu}$.

4. Since for $t \in [\![T_G, 1]\!]$ we have that $\mathbf{P}_G \not\ll \mathbf{P}_t^\omega$ and $\mathbf{P}_G \perp \mathbf{P}_t^\omega$ if $\sigma(G)$ is non-atomic it follows from (1.85) that equivalently $\mathbf{P} \not\ll \mathbf{P}^z$ and $\mathbf{P}^z \perp \mathbf{P}$ on $\mathcal{F}_t$ for $t \in [\![T_G, 1]\!]$ if $\sigma(G)$ is non-atomic.

On Wiener space Imkeller (1996) has recently shown how to get explicit expressions for the compensator using the Clark-Ocone representation formula. The following lemma presents his result and summarizes complementary results from Jacod (1980) .

Lemma 20 *If "condition A" and (i) $p(\omega, s, z) \in \mathbb{L}^{1,2}(\mathbb{R}^d)$ as well as (ii) the mapping $r \mapsto \mathcal{D}_r p(\omega, s, z)$ is left-continuous in $L^1(\Omega)$ at $s \in [0, t]$ and $z \in \mathbb{R}^q$ are satisfied then we have for $t < \tau^z$ that*

$$p(\omega, t, z) = q(z)\mathcal{E}(\int_0^\cdot (\alpha_s^z)^* d\omega(s))_t \tag{1.89}$$

where the stopping time τ^z is given by

$$\tau^z := \inf\{u \in [0, 1] : p(\omega, u, z) = 0\}, \tag{1.90}$$

whereas

$$(\alpha_t^z(\omega))^* := \frac{\mathcal{D}_t p(\omega,t,z)}{p(\omega,t,z)}. \tag{1.91}$$

Furthermore for all $t \in [\![0, T_G[\![$ *we have that* $p(\omega,t,G) > 0$ $\mathbf{P}$- *a.s. and therefore*

$$(\alpha_t^G(\omega))^* = \left[\frac{\mathcal{D}_t p(\omega,t,z)}{p(\omega,t,z)}\right]_{z=G} \tag{1.92}$$

is well defined for all $t \in [\![0, T_G[\![$.

Proof

Since under "condition A" $\mathbf{P}_t^\omega(dz) = p(\omega,t,z)\boldsymbol{\nu}(dz)$ for $t \in [\![0, T_G[\![$ and therefore

$$\mathbf{E}[\mathbf{P}_t^\omega(B)|\mathcal{F}_s] = \mathbf{E}[\int_B p(\omega,t,z)\boldsymbol{\nu}(dz)|\mathcal{F}_s] \tag{1.93}$$

and by Fubini's theorem

$$\mathbf{E}[\int_B p(\omega,t,z)\boldsymbol{\nu}(dz)|\mathcal{F}_s] = \int_B \mathbf{E}[p(\omega,t,z)|\mathcal{F}_s]\boldsymbol{\nu}(dz) \tag{1.94}$$

But since at the same time

$$\mathbf{E}[\mathbf{P}_t^\omega(B)|\mathcal{F}_s] = \mathbf{P}_s^\omega(B) = \int_B p(\omega,s,z)\boldsymbol{\nu}(dz) \tag{1.95}$$

we have $\mathbf{P}(d\omega) \otimes \boldsymbol{\nu}(dz)$ a.e. that $\mathbf{E}[p(\omega,t,z)|\mathcal{F}_s] = p(\omega,s,z)$ and we have established that the conditional density process $p(\omega,t,z)$ is a non-negative martingale.

Since $p(\omega,s,z) \in \mathbb{L}^{1,2}(\mathbb{R}^d)$ we can use the Clark-Ocone formula (see Nualart (1995) proposition 1.3.5. page 42) to represent the conditional density as

$$p(\omega,t,z) = q(z) + \int_0^t \mathbf{E}[\mathcal{D}_v p(\cdot,t,z)|\mathcal{F}_v]d\omega(v) \tag{1.96}$$

and since the conditional expectation operator and the Malliavin derivative commute (see Nualart (1995) proposition 1.2.4 page 32) we have for all $r \in [0,v]$ that

$$\mathbf{E}[\mathcal{D}_r p(\cdot,t,z)|\mathcal{F}_v] = \mathcal{D}_r \mathbf{E}[p(\cdot,t,z)|\mathcal{F}_v] = \mathcal{D}_r p(\omega,v,z) \tag{1.97}$$

Then since $\mathcal{D}_r p(\omega, v, z)$ is assumed to be left-continuous $\lim_{r\uparrow v} \mathcal{D}_r p(\omega, v, z) = \mathcal{D}_v p(\omega, v, z)$ exists in $L^1(\Omega)$ and we can write

$$p(\omega, t, z) = q(z) + \int_0^t \mathcal{D}_v p(\cdot, v, z) d\omega(v) \tag{1.98}$$

We clearly have that $p(\omega, t, z) > 0$ for $t < \tau^z$. Suppose that for some $t \geq \tau^z$ we also have that $p(\omega, t, z) > 0$. Since the conditional density process is a martingale and τ^z is a $\mathcal{F}_t$- stopping time, we would have that $p(\omega, \tau^z, z) > 0$ since by the optional sampling theorem $p(\omega, \tau^z, z) = \mathbf{E}[p(\cdot, t, z)|\mathcal{F}_{\tau^z}]$, a contradiction since obviously $p(\omega, \tau^z, z) = 0$.We therefore have shown that $p(\omega, t, z) > 0$ for $t < \tau^z$ and $p(\omega, t, z) = 0$ for $t \geq \tau^z$.

It follows for $t < \tau^z$ that

$$p(\omega, t, z) = q(z) + \int_0^t p(\cdot, v, z) \frac{\mathcal{D}_v p(\cdot, v, z)}{p(\cdot, v, z)} d\omega(v) \tag{1.99}$$

or equivalently that

$$p(\omega, t, z) = q(z)\mathcal{E}(\int_0^{\cdot} \frac{\mathcal{D}_v p(\cdot, v, z)}{p(\cdot, v, z)} d\omega(v))_t \tag{1.100}$$

This establishes (1.91).

To show that for all $t \in [\![0, T_G[\![$ we $\mathbf{P}$- a.s. have that $p(\omega, t, G) > 0$ is equivalent to show that $\mathbf{P}(\tau^G = 1) = 1$. Clearly for all $t \in [0, 1[$

$$\mathbf{E}[\int_{R^q} p(\cdot, t, z)\boldsymbol{\nu}(dz)] = 1 \tag{1.101}$$

But since $\mathbf{P}$- a.s.

$$\int_{R^q} p(\omega, t, z)\boldsymbol{\nu}(dz) = \int_{R^q} \mathbf{1}_{\{\tau^z > t\}} p(\omega, t, z)\boldsymbol{\nu}(dz) \tag{1.102}$$

and

$$\mathbf{E}[\int_{R^q} \mathbf{1}_{\{\tau^z > t\}} p(\cdot, t, z)\boldsymbol{\nu}(dz)] = \mathbf{E}[\mathbf{1}_{\{\tau^G > t\}}] \tag{1.103}$$

we have established for all $t \in [\![0, T_G[\![$ that $\mathbf{P}(\tau^G > t) = 1$ and therefore $\tau^G = T_G$ $\mathbf{P}$- a.s..

Q.E.D.

Remarks

1. Since for non-atomic signals we have for $t \in [\![T_G, 1]\!]$ that $\mathbf{P}^z \perp \mathbf{P}$ on $\mathcal{F}_t$ it follows for all $z \in \mathbb{R}^q$ that

$$\lim_{t \uparrow T_G} \mathcal{E}\left(\int_0^{\cdot} (\alpha_s^z)^* d\omega(s)\right)_t = \begin{cases} +\infty & \mathbf{P}^z\text{- a.s.} \\ 0 & \mathbf{P}\text{- a.s.} \end{cases} \tag{1.104}$$

   whereas for atomic signals

$$\lim_{t \uparrow T_G} \mathcal{E}\left(\int_0^{\cdot} (\alpha_s^z)^* d\omega(s)\right)_t = \frac{1}{\mathbf{P}_G(\{z\})} \tag{1.105}$$

   $\mathbf{P}^z$- a.s. and

$$\mathbf{P}\left(\left\{\lim_{t \uparrow T_G} \mathcal{E}\left(\int_0^{\cdot} (\alpha_s^z)^* d\omega(s)\right)_t = +\infty\right\}\right) > 0 \tag{1.106}$$

   and therefore $\mathbf{P}^z \ll \mathbf{P}$ on $\mathcal{F}_t$ for all $t \in [0, 1]$ in this case.

2. For signals which have laws that are absolutely continuous with respect to the Lebesgue measure Imkeller's (1996) theorem 5 gives for signals absolutely continuous with respect to the Lebesgue measure sufficient conditions for which (i) and (ii) are satisfied.

From the law of W given $G = z$ we can find the $(\mathbf{P}^z, \mathbb{F})$- compensator of the Wiener process W by a Girsanov transformation. Since

$$\frac{\mathbf{E}[\mathcal{E}(\int_0^{\cdot} (\alpha_s^z)^* d\omega(s))_{t \wedge \tau^z}]}{q(z)} = 1 \tag{1.107}$$

it follows that for $t < \tau^z$

$$W(\omega)_t^z := \omega(t) - \int_0^t \alpha_s^z(\omega) ds \tag{1.108}$$

is a $(\mathbf{P}^z, \mathbb{F})$- Wiener process for all $z \in \mathbb{R}^q$. Furthermore if $\mathbf{P}^z$- a.s. we have that

$$\int_0^{\tau^z} \alpha_s^z ds < +\infty \tag{1.109}$$

then W^z can be extended for $t \geq \tau^z$ as follows

$$W(\omega)_t^z := \omega(t) - \int_0^{t \wedge \tau^z} \alpha_s^z(\omega)^* ds \tag{1.110}$$

An investor "gambling" upon the event $\{G = z\}$ has states of nature given by W^z. Since an insider can be seen as a "gambler" who knows with certainty the realization of the event he is "gambling" upon we get states for insiders by evaluating "gamblers" states at $z = G$. That W^G on $[\![0, T_G[\![$ are indeed the state of natures for an insider follows from theorem 1 of Foellmer and Imkeller (1993) . They show that W^G corresponds to the $(\mathbf{P}, \mathbb{G})$- decomposition of the Wiener process W. Since we want to analyze hedging and investment policies for investment horizons longer than the resolution time of the signal without loosing the semi-martingale property of processes relevant for insiders' decisions we have to assume the following

Assumption 6 *Signals G are such that*

$$\int_0^{T_G} \alpha_s^G ds < \infty \tag{1.111}$$

$\mathbf{P}$-*a.s.*

Remarks

1. An example where assumption 6 is satisfied is $G = \omega(T)$ some $T \in [0, 1]$. In this case the resolution time is $T_G = T$ and

$$\int_0^{T_G} \alpha_s^G ds = \int_0^{T_G} \log(T_G - s) d\omega(s) \tag{1.112}$$

P-a.s. such that

$$W_t^G = \mathbf{E}[\int_0^{T_G} (T_G + \log(T_G - s))d\omega(s)|\mathcal{G}_t] \quad (1.113)$$

which proves that W^G is an uniformly integrable $(\mathbf{P}, \mathbb{G})$- martingale. It follows by arguments similar to those in Jeulin and Yor (1977) that W_T^G is independent of $\omega(T)$ and that $\mathbb{G}$ corresponds for $t \in [0, T]$ to the filtration of the Brownian bridge $\beta_t(\omega) := \omega(t) - \frac{t}{T}G(\omega)$.

2. In the paper of Elliott,Geman and Korkie (1997) (1.111) is satisfied, since they introduce insider information on incomplete information such that non-insiders never get to know the signal. Consequently $T_G = 1$ and since for their signal "condition A" holds for all $t \in [0, 1]$ assumption 6 is satisfied for any investment horizon in their model.

Theorem 21 *If "condition A" and (i) and (ii) of lemma 20 and assumption 6 are satisfied then the process* $W^G = (W_t^G, t \in [0, 1])$ *given by*

$$W^{G(\omega)}(\omega)_t = \omega(t) - \int_0^{t \wedge T_G} (\alpha_s^{G(\omega)}(\omega)))^* ds \quad (1.114)$$

is a $(\mathbf{P}, \mathbb{G})$- *Wiener process.*

Proof

Obviously W^G is $\mathcal{G}_t$ adapted. Since for all $A = A_1 \times A_2 \in \mathcal{F}_s \otimes \sigma(G)$ where $s \in [0, t]$ and $t < T_G$

$$\mathbf{E}[\mathbf{1}_A \mathbf{E}[W_t^G - W_s^G|\mathcal{G}_s]] = \int_{\Omega \times R^q} \mathbf{1}_{A_1 \times A_2}(w, z)(W_t^z - W_s^z)\mathbf{P}^z(d\omega)\mathbf{P}_G(dz) \quad (1.115)$$

it follows that

$$\mathbf{E}[\mathbf{1}_A \mathbf{E}[W_t^G - W_s^G|\mathcal{G}_s]] = \int_{R^q} \mathbf{1}_{A_2}(z)\mathbf{E}^{\mathbf{P}^z}[\mathbf{1}_{A_2}\mathbf{E}^{\mathbf{P}^z}[W_t^z - W_s^z|\mathcal{F}_s]]\mathbf{P}_G(dz) \quad (1.116)$$

From Girsanov's theorem we know that W^z is a $\mathbb{F}$- Wiener process. It follows that $\mathbf{E}^{\mathbf{P}^z}[W_t^z - W_s^z|\mathcal{F}_s] = 0$ for $s < t$ and therefore that

$$\mathbf{E}[\mathbf{1}_A \mathbf{E}[W_t^G - W_s^G|\mathcal{G}_s] = 0 \tag{1.117}$$

The statement that W^G is a $(\mathbf{P}, \mathbb{G})$- Wiener process on $[\![0, T_G[\![$ follows then from Lévy's characterization of Brownian motion since the quadratic variation of W^G is

$$[(W^G)^i, (W^G)^j]_t = \delta_{i,j} t \tag{1.118}$$

for all $i, j \in \{1, \ldots, d\}$.

For $t \geq T_G$ we have $\mathcal{G}_t = \mathcal{F}_t$ and $dW_t^G = d\omega(t)$. It follows that W^G corresponds to a Wiener process starting at $-\int_0^{T_G} \alpha_s^G ds$ on $[\![T_G, 1]\!]$.

Q.E.D.

Remarks

1. Given the expression for the process W^G on $[\![0, T_G[\![$ we might from Girsanov's theorem conclude that the measure determined by the density process $\mathcal{E}(-\int_0^{\cdot}(\alpha_s^G)^* dW_s^G)$. should correspond to the Wiener measure $\mathbf{P}$ and therefore be constant. Imkeller and Foellmer (1993) have shown that in fact this density process with respect to the Wiener measure defines a measure which is not even equivalent to the Wiener measure, a paradox, which is explained by the fact that $\frac{d\mathbf{P}}{d\mathbf{P}^G}$ corresponds on $\mathcal{G}_t$ for $t \in [0, 1]$ to the Radon-Nikodym derivative of the product measure $\mathbf{P} \otimes \mathbf{P}_G$ with respect to the joint law $\mathbf{R}$, that is defines a product measure on $\sigma(G) \otimes \mathcal{F}_t$ which corresponds to the Wiener measure only if projected on its second coordinate. It is this fact which is crucial for whether or not on insider information local martingale measures exist.

2. If we compare the process W^G and W on $[\![T_G, 1]\!]$ we see that they have the same increments but not the same starting point $W_{T_G} - W^G_{T_G} = \int_0^{T_G} \alpha^G_s ds$.

3. Since any $\mathcal{F}_t$- local martingale is given as $\int_0^{\cdot} \phi^*_s d\omega(s)$ its $\mathcal{G}_t$- compensator is given by $\int_0^{\cdot \wedge T_G} \phi^*_s \alpha^G_s ds$ as long this integral is finite $\mathbf{P}$- a.s..

1.8 Appendix B: The Representation of Contingent Claims as Two-Parameter Random Variables

The results of appendix 1.7 show how using Girsanov transformations the Doob-Meyer decompositions of $(\mathbf{P},\mathbb{F})$- local martingales can be obtained on $(\Omega,\mathcal{F}_1,\mathbb{G},\mathbf{P})$ in two steps. First conditional on an arbitrary realization of the signal we get $(\mathbf{P}^z,\mathbb{F})$- semi-martingales by a change of measure. Secondly we get desired representations for enlarged filtration simply be evaluating this decomposition at the realization of the signal. This shows that processes on insiders' stochastic basis can be interpreted as two-parameter processes, one parameter for the state of nature and the other for the realization of the signal. We will show how this can be used to get optimal hedging strategies for insiders. To do this we need to establish that $\mathcal{G}_T$- measurable (some $T<T_G$) contingent claim can be regarded as $\mathcal{F}_T\otimes\sigma(G)$ measurable mappings from $\Omega\times\mathbb{R}^q$ to $\mathbb{R}^l$ some $l\in\mathbb{N}$. The next theorem will show that this is possible for sufficiently smooth signals.

Theorem 22 *Under "condition A", (i) and (ii) of lemma 20 and assumption 6 we have for any $\mathcal{G}_T$- measurable random variable H such that $H\in\mathbb{D}^{1,1}(\mathbb{R}^q)$ if*

$$\int_0^{T_G}|\mathbf{E}[\mathcal{D}_tH|\mathcal{F}_t]\alpha_t^G|dt<+\infty \qquad (1.119)$$

$\mathbf{P}$- *a.s. that*

$$H(\omega)=C^{G(\omega)}(\omega) \qquad (1.120)$$

where for all $z\in\mathbb{R}^q$

$$C^z(\omega)=\mathbf{E}[H|\mathcal{F}_T](\omega)+\mathbf{E}^{\mathbf{P}^z}[\int_T^{T_G}\mathbf{E}[\mathcal{D}_tH|\mathcal{F}_t]\alpha_t^z dt|\mathcal{F}_T](\omega) \qquad (1.121)$$

$\mathbf{P}$- *a.s.*.

In the proof and for other results we need the following lemma.

Lemma 23 *For $\mathcal{G}_{T_G-}$- measurable random variables $H(\omega)$ that can be written as $H(\omega) = C^{G(\omega)}(\omega)$ where $C^z(\omega)$ is $\mathbf{F}_{T_G-} \otimes \sigma(G)$- measurable we have under "condition A" for all $t \in [0, T_G[$ that*

$$\left(\mathbf{E}^{\mathbf{P}^z}[C^z|\mathcal{F}_t]\right)_{|z=G} = \mathbf{E}[C^G|\mathcal{G}_t] \tag{1.122}$$

whereas

$$(\mathbf{E}[C^z|\mathcal{F}_t])_{|z=G} = \mathbf{E}[\frac{p(\cdot,t,G)}{p(\cdot,T,G)} H|\mathcal{G}_t] \tag{1.123}$$

Proof

Since $\mathbf{1}_E(\omega) = \mathbf{1}_{G(E)}(G(\omega))$ we have for all $E \in \mathcal{G}_t$ that

$$\mathbf{E}[\mathbf{1}_E \mathbf{E}[C^G|\mathcal{G}_t]] = \mathbf{E}[\mathbf{1}_{G(E)}(G)C^G] \tag{1.124}$$

and since $\mathbf{E}[\mathbf{1}_{G(E)}(G)C^G] = \int_{G(E)\times\Omega} C^z(\omega)\mathbf{R}(dz, d\omega)$ that

$$\mathbf{E}[\mathbf{1}_E \mathbf{E}[C^G|\mathcal{G}_t]] = \int_{G(E)\times\Omega} C^z(\omega)\mathbf{R}(dz, d\omega) \tag{1.125}$$

Then since at the same time

$$\mathbf{E}[\mathbf{1}_E \left(\mathbf{E}^{\mathbf{P}^z}[C^z|\mathcal{F}_t]\right)_{|z=G}] = \int_{G(E)} \mathbf{E}^{\mathbf{P}^z}[C^z|\mathcal{F}_t]\mathbf{P}_G(dz) \tag{1.126}$$

and

$$\int_{G(E)} \mathbf{E}^{\mathbf{P}^z}[C^z|\mathcal{F}_t]\mathbf{P}_G(dz) = \int_{G(E)\times\Omega} C^z(\omega)\mathbf{R}(dz, dw) \tag{1.127}$$

we have shown (1.122). Then since $\mathbf{1}_{\{T<\tau^G\}} = 1$ and for $T < \tau^z$ by Bayes' law $\mathbf{E}^{\mathbf{P}^z}[C^z|\mathcal{F}_t] = \mathbf{E}[\frac{p(\cdot,t,z)}{p(\cdot,T,z)} C^z|\mathcal{F}_t]$ we obtain (1.123) by repeating the same argument for $\tilde{C}^z(\omega) = \frac{p(\omega,t,z)}{p(\omega,T,z)} \mathbf{1}_{\{T<\tau^z\}} C^z(\omega)$.

Q.E.D.

Proof of theorem 22

By assumption $H \in \mathbb{D}^{1,1}(\mathbb{R}^d)$ and therefore from the Clark-Ocone representation formula we have since H is $\mathcal{F}_{T_G}$- measurable that

$$H = \mathbf{E}[H] + \int_0^{T_G} \mathbf{E}[\mathcal{D}_u H|\mathcal{F}_u] d\omega(u) \tag{1.128}$$

Since for all $d\omega(t) = dW_t^z - \alpha_t^z du$ and H is by assumption $\mathcal{G}_T$- measurable we get

$$H = \mathbf{E}[H|\mathcal{F}_T] + \mathbf{E}[\int_T^{T_G} \mathbf{E}[\mathcal{D}_u H|\mathcal{F}_u]\alpha_u^G du|\mathcal{G}_T] \tag{1.129}$$

where the second integral is well defined since (1.119) holds. Since from lemma 23 we know that

$$\left(\mathbf{E}^{\mathbf{P}^z}[\int_T^{T_G} \mathbf{E}[\mathcal{D}_t H|\mathcal{F}_t]\alpha_t^z dt|\mathcal{F}_T]\right)_{|z=G} = \mathbf{E}[\int_T^{T_G} \mathbf{E}[\mathcal{D}_u H|\mathcal{F}_u]\alpha_u^G du|\mathcal{G}_T] \tag{1.130}$$

we have established that (1.120) must hold with (1.121).

Q.E.D.

Remarks

1. Since $\mathcal{D}_t H = 0$ for $t \in [T, 1]$ whenever the contingent claim H is also $\mathcal{F}_T$-measurable such claims do not depend on the signal.
2. Theorem 22 shows that whenever a claim is only measurable on insider information, it can be regarded as $\mathcal{F}_T \otimes \sigma(G)$- measurable random variable.

1.9 Appendix C: The Domain of Malliavin's Derivative and its Adjoint

The domain $\mathbb{L}^{1,2}(\mathbb{R}^d)$ of the Skorohod integral, that is the adjoint operator of the Malliavin derivative or divergence operator, corresponds to the Hilbert space with norm

$$\| u \|_{L^{1,2}(R^d)} := (\mathbf{E}[\int_0^1 u_s^2 ds])^{1/2} + (\mathbf{E}[\int_0^1 ds \int_0^1 \mathcal{D}_t u_s (\mathcal{D}_t u_s)^* dt])^{1/2} \tag{1.131}$$

The operator $\mathcal{D}_t$ in the Hilbert norm above denotes the Malliavin derivative (or divergence) operator with domain $\mathbb{D}^{1,p}(\mathbb{R}^d)$. The domain corresponds to the Banach space given by the completion of the set of smooth random variables $F \in \mathcal{S}$ on $(\Omega, \mathcal{F}_1, \mathbf{P})$, i.e. random variables of the form

$$F = f(\omega(t_1), \dots, \omega(t_m)), \text{ some } t_1, \dots t_m \in [0,1], f \in C_0^\infty(\mathbb{R}^{m \times d}) \tag{1.132}$$

with respect to the norm

$$\| F \|_{1,p} := (\mathbf{E}[F^p])^{1/p} + (\mathbf{E}[\int_0^1 \|\mathcal{D}_t F\|^p dt])^{1/p} \tag{1.133}$$

where for smooth random variables $F \in \mathcal{S}$ Malliavin derivatives are given by

$$\mathcal{D}_t F := \sum_{i=1}^d \frac{\partial}{\partial x_i} f(\omega(t_1), \dots \omega(t_m)), \quad t \in [0,1] \tag{1.134}$$

For definitions concerning Malliavin Calculus we refer to Nualart (1995) .

1.10 Appendix D: Proofs

1.10.1 Proofs of section 1.3

Proof of theorem 1

By assumption 3 the Doob-Meyer decomposition of W on $(\Omega, \mathcal{F}_1, \mathbb{G}, \mathbf{P})$ is as shown in Appendix 1.7 given for $t \in [0,1]$ by

$$\omega(t) = W_t^G + \int_0^{t \wedge T_G} \alpha_s^G ds \tag{1.135}$$

Replacing this in the representation on $(\Omega, \mathcal{F}_1, \mathbb{F}, \mathbf{P})$ gives the desired result for prices, dividend rate and endowment rate processes since by assumption corresponding volatility coefficients are bounded and assumption 3 holds. Similarly the additional condition imposed on portfolio policies gives the decomposition of wealth.

Q.E.D.

Proof of corollary 2

Clearly if $t \in [\![0, T_G[\![$ and $\mathcal{F}_t \perp \sigma(G)$ we have for all $t \in [\![0, T_G[\![$ that $p(\omega, t, z) = q(z)$ and therefore $\alpha_t^z = 0$. That independence is also necessary follows from the fact that Malliavin derivatives are zero for all $t \in [\![0, T_G[\![$ if and only if $p(\omega, t, z)$ lies in the subspace of the zeroth Wiener chaos and is therefore constant (Nualart (1995) page 31). It follows that $\alpha_t^z = 0$ implies that G is independent of $\mathcal{F}_t$.

Q.E.D.

Proof of proposition 6

If we define $E := \{\frac{X_T^{\pi,e}}{e_0 B_T} \geq 1\}$ and $F := \{\frac{X_T^{\pi,e}}{e_0 B_T} > 1\}$ we have since $\mathbf{P}(E) = \mathbf{E}[\mathbf{P}(E|\mathcal{G}_0)]$ and $\mathbf{P}(F) = \mathbf{E}[\mathbf{P}(F|\mathcal{G}_0)]$ that there is an arbitrage whenever there is a conditional arbitrage.

It is possible that an arbitrage has an associated gains from trade of the form $\alpha_t \mathbf{1}_E$ some $E \in \mathcal{G}_0$ where $\mathbf{P}(E) > 0$ and $\mathcal{G}_t$- adapted process α. Since on E^c gains from trade are zero there is no arbitrage conditional on E^c. This shows that the existence of an arbitrage does not imply the existence of a conditional arbitrage.

Q.E.D.

Proof of theorem 8

Since for $S_t^G := -\int_0^T r_s ds - \int_0^T (\theta_s + \alpha_s^G)^\star dW_s^G$ we have that

$$\left(\mathcal{E}(S^G)_t B_t\right)^2 \exp\{-\int_0^t \|\theta_s + \alpha_s^G\|^2 ds\} = \mathcal{E}(2S^G)_t B_t^2 \tag{1.136}$$

it follows that

$$\{\mathcal{E}(2S^G)_t B_t^2 > 0\} = \{\mathcal{E}(S^G)_t B_t > 0\} \bigcap \{\int_0^t \|\theta_s + \alpha_s^G\|^2 ds < \infty\} \tag{1.137}$$

Then since $\mathcal{E}(2S^G)_t B_t^2 > 0$ if and only if $\mathcal{E}(S^G)_t B_t > 0$ it follows that

$$\{\int_0^t \|\theta_s + \alpha_s^G\|^2 ds = +\infty\} = \{\mathcal{E}(\int_0^\cdot (\theta_s + \alpha_s^G)^\star dW_s^G)_t = 0\} \tag{1.138}$$

Since the market price of risk on public information is by the assumptions on the spot rate, the appreciation rates and volatilities of stocks bounded and from the representation of the density process between the conditional and unconditional Wiener measure for any $t \in [\![0, T_G[\![$ it must hold true that $\int_0^t \|\alpha_s^G\|^2 ds < +\infty$, we have for any admissible sequence of portfolio policies π^n and positive sequence $(\delta^n)_{\in \mathbf{N}}$ such that $\frac{X^{\pi^n,e}_t}{e_0 B_t} > \delta^n$ for all $t \in [0,1]$ that

$$\mathcal{E}(S^G)_{t\wedge T_G} \frac{X^{\pi,e}_{t\wedge T_G} - \delta^n B_{t\wedge T_G}}{e_0 - \delta^n} = \mathcal{E}(\int_0^{\cdot} (\frac{\pi^\star_s \sigma_s}{X^{\pi,e}_s - \delta^n B_s} - (\theta_s + \alpha^G_s)^\star) dW^G_s)_{t\wedge T_G} \tag{1.139}$$

First we consider an investment horizon which ends before resolution time $T < T_G$. In this case we have that $\mathbf{P}^z(\{\frac{d\mathbf{P}}{d\mathbf{P}^z}_{|\mathcal{F}_T} \in]0, +\infty[\}) = 1$ and therefore

$$\mathbf{E}^{\mathbf{P}^z}[\mathcal{E}(S^z_T B_T] = \mathbf{E}[\mathcal{E}(-\int_0^{\cdot} \theta^\star_s dW_s)_T B_T] \tag{1.140}$$

But since by the boundedness of the coefficients of price processes for common information Novikov's condition

$$\mathbf{E}[\exp\{\frac{1}{2}\int_0^T \|\theta_s\|^2 ds\}] < +\infty \tag{1.141}$$

is satisfied and therefore it follows by proposition 1.15 p.308 of Revuz and Yor (1990) that $\mathcal{E}(S_t)B_t$ is an uniformly integrable martingale. But this implies that $\mathbf{E}[\mathcal{E}(S_T)B_T] = 1$. Consequently since $\mathbf{P}^z \sim \mathbf{P}$ on $\mathcal{F}_T$ whenever $T < T_G$ we must have that

$$\mathbf{E}^{\mathbf{P}^z}[\mathcal{E}(S^z)_T B_T] = 1 \tag{1.142}$$

or equivalently

$$\mathbf{E}[\mathcal{E}(S^G)_T B_T | \mathcal{G}_0] = 1 \tag{1.143}$$

This shows that the measure $d\tilde{\mathbf{Q}} := \mathcal{E}(S^G_T) B_T d\mathbf{P}^z_{|z=G}$ is a local martingale measure for the insider.

Since the left hand side of (1.139) is non-negative and a local martingale, it must be a super-martingale and therefore

$$\mathbf{E}^{\mathbf{P}^z}[\mathcal{E}(S^z)_T(\frac{X^{\pi,e}_T - \delta^n B_T}{e_0 - \delta^n} - \delta^n B_T)] \leq 1 - \delta^n \tag{1.144}$$

Consequently it must from Markov's inequality hold true that for any sequence of portfolio strategies such that $\frac{X_T^{\pi^n,e}}{e_0 B_T} > \delta^n$ and $\epsilon > 0$ that

$$\mathbf{E}[\mathcal{E}(S^G)_T B_T \mathbf{1}_{\{|\frac{X_T^{\pi,e}-\delta^n B_T}{e_0-\delta^n}-\delta^n B_T|>\epsilon\}}|\mathcal{G}_0] \le \frac{1}{\epsilon}(1-\delta^n) \tag{1.145}$$

such that whenever $\delta^n \to 1$ for $n \to +\infty$ we must since $\mathbf{P}^z(\{\mathcal{E}(S^z)_T B_T > 0\}) = 1$ have that

$$\lim_{n\to+\infty} \mathbf{P}(|\frac{X_T^{\pi^n,e}}{e_0 B_T} - 1| > \epsilon|\mathcal{G}_0) = 0 \tag{1.146}$$

where we have used that $\frac{X_T^{\pi^n,e}}{e_0 B_T} - 1$ must converge in probability to zero whenever $\frac{X_T^{\pi,e}-\delta^n B_T}{e_0-\delta^n} - \delta^n B_T$ does. It follows thats there are no free lunches with vanishing risk for an investment horizon shorter than resolution time.

Next we show that it is necessary for NFLVR for an investment horizon T that $\mathbf{P}(\{\mathcal{E}(S^G)_T B_T > 0\}|\mathcal{G}_0) = 1$. Clearly if $\mathbf{P}(\{\mathcal{E}(S^G)_T B_T = 0\}|\mathcal{G}_0) > 0$ we must have that $\mathbf{P}^z(\{\frac{1}{\mathcal{E}(S^z)_T B_T} = +\infty\}) > 0$ at $z = G$. But then since the sequence of mean-variance strategies

$$\tilde{\pi}_t^n := (\sigma_t \sigma_t^\star)^{-1}(b_t + \sigma_t \alpha_t^G - 1_d r_t)(X_t - \delta^n B_t)\mathbf{1}_{[\![0,T_G[\![}(t) \tag{1.147}$$

are admissible and have an associated wealth process such that

$$\mathcal{E}(S^G)_T B_T \frac{X_T^{\tilde{\pi},e} - \delta^n B_T}{e_0 - \delta^n B_T} = \frac{\mathcal{E}(S)_T B_T}{\mathcal{E}(S)_{T_G} B_{T_G}} \tag{1.148}$$

We therefore would have that

$$\mathbf{P}^z(\frac{X_T^{\tilde{\pi},e} - \delta^n B_T}{e_0 - \delta^n B_T} = +\infty) > 0 \tag{1.149}$$

which is impossible if for all $\epsilon > 0$ we have that $\lim_{n\to+\infty} \mathbf{P}(|\frac{X_T^{\tilde{\pi}^n,e}}{e_0 B_T}| > \epsilon) = 0$, which must hold true because there is no free lunch with vanishing risk. It follows that $\mathbf{P}^z(\{\mathcal{E}(S^z)_T B_T = 0\}) = 0$ whenever NFLVR is satisfied.

Since $\mathbf{P} \not\ll \mathbf{P}^z$ on $\mathcal{F}_{T_G}$ and therefore $\mathbf{P}^z(\{\frac{d\mathbf{P}}{d\mathbf{P}^z}|_{\mathcal{F}_{T_G}} = 0\}) > 0$ and since $\mathbf{P} \sim \mathcal{E}(S)_{T_G} B_{T_G} \cdot \mathbf{P}$ we must have that

$$\mathbf{P}(\{\mathcal{E}(S^G)_{T_G} B_{T_G} = 0\}|\mathcal{G}_0) > 0 \tag{1.150}$$

and consequently that, whenever an insider has an investment horizon longer than the resolution time he will also have a free lunch with vanishing risk given by the sequence of mean-variance strategies given in (1.147).

It remains to be shown that for $T \geq T_G$ there is an arbitrage if an only if $\mathcal{G}_0$ is non-atomic.

If we put for all $n \in \mathbb{N}$ $\delta^n = K$ we obtain for tame portfolios from (1.139) that

$$\mathcal{E}(S^G)_{t\wedge T_G}\frac{X^{\pi,e}_{t\wedge T_G} - KB_{t\wedge T_G}}{e_0 - K} = \mathcal{E}(\int_0^{\cdot}(\frac{\pi^\star_s\sigma_s}{X^{\pi,e}_s - KB_s} - (\theta_s + \alpha^G_s)^\star)dW^G_s)_{t\wedge T_G} \tag{1.151}$$

Then suppose that $\mathcal{G}_0$ is non-atomic. As we have seen in this case we have that $\mathbf{P}^z \ll \mathbf{P}$ on $\mathcal{F}_T$ even for $T > T_G$. Consequently

$$\mathbf{E}^{\mathbf{P}^z}[\mathcal{E}(S^z)_T B_T] = \mathbf{E}[\mathbf{1}_{\{\frac{dP}{dP^z}|\mathcal{F}_T > 0\}}\mathcal{E}(S)_T B_T] \tag{1.152}$$

But since $P(\{\frac{d\mathbf{P}^z}{d\mathbf{P}} = +\infty\}) = 0$ and that from Novikov's condition we have already argued that $\mathcal{E}(S)_t B_T$ is an uniformly integrable martingale starting at 1, it follows that

$$\mathbf{E}^{\mathbf{P}^z}[\mathcal{E}(S^z)_T B_T] = 1 \tag{1.153}$$

Then since by the same argument as before the left hand side of (1.151) is non-negative local martingale and therefore a super-martingale starting at 1, it must hold true for any admissible portfolio strategy that

$$\mathbf{E}^{\mathbf{P}}[\mathcal{E}(S)_T B_T(\frac{\frac{X^{\pi,n}_T}{B_T} - K}{e_0 - K} - 1)] \leq 1 \tag{1.154}$$

Since $\mathcal{E}(S)_T B_T > 0$ $\mathbf{P}$- a.s. we must have that $\mathbf{P}(\{\frac{X^{\pi,e}_T}{e_0 B_T} > 1\}) = 0$ in this case. Since $\mathbf{P}^z \ll \mathbf{P}$ on $\mathcal{F}_T$ for all $T \in [0,1]$ it must therefore also hold true that

$$\mathbf{P}(\{\frac{X^{\pi,e}_T}{e_0 B_T} > 1\}|\mathcal{G}_0) = 0 \tag{1.155}$$

That is there is no arbitrage for the insider when his anticipative signal is atomic.

Finally if $\mathcal{G}_0$ is non-atomic we have that for any that the mean variance strategy

$$\tilde{\pi}_t = (\sigma_t \sigma_t^\star)^{-1}(b_t + \sigma_T \alpha_t^G - 1_d r_t)(X_t - B_t K)\mathbf{1}_{]T_G, T_G]}(t) \tag{1.156}$$

has an associated wealth process such that

$$\frac{X_T^{\tilde{\pi},e} - KB_T}{e_0 - K} = \frac{\mathcal{E}(S^G)_{T_G-}}{\mathcal{E}(S^G)_{T_G}} \tag{1.157}$$

But since on $\mathcal{F}_{T_G}$ we have that $\mathbf{P}^z \perp \mathbf{P}$ and consequently $\frac{d\mathbf{P}}{d\mathbf{P}^z}_{|\mathcal{F}_{T_G}} = 0$ $\mathbf{P}^z$- a.s. we have at $z = G$ that $\frac{X_T^{\tilde{\pi},e} - KB_T}{e_0 - K} = +\infty$ $\mathbf{P}^z$- a.s. and therefore also that

$$\mathbf{P}(\frac{X_T^{\tilde{\pi},e}}{e_0 B_T} > 1 | \mathcal{G}_T) = 1 \tag{1.158}$$

and therefore $\tilde{\pi}$ is an arbitrage.

Q.E.D.

Proof of corollary 9

For signals $G = G^0 + Z$ where Z is independent from G^0 we have that $\sigma(G) = \sigma(G^0) \bigvee \sigma(Z)$. And consequently that for $t \geq T_{G^0}$ that $\mathbf{P}(G \in B|\mathcal{F}_t) = \mathbf{P}(Z \in (B - x))_{|x=G} > 0$ for some $B \in \mathcal{B}_{\mathbf{R}^d}$, such that $T_G = 1$ $\mathbf{P}$- a.s.. It follows from theorem 8, that there are no arbitrage opportunities even for non-atomic insider information.

Q.E.D.

1.10.2 Proofs of section 1.4

Proof of theorem 10

For the contingent claim H define the tracking error $\phi_T^{\pi^z,c^z,z} := (X_T^{\pi^z,c^z} - C^z)$ where C^z is given by (1.121). Then since by assumption $\mathcal{E}(S)_T C^z \in \mathbb{D}^{1,1}(\mathbb{R}^d)$ we have from the Clark-Ocone formula and the fact that $\mathcal{E}(S)_T C^z$ is $\mathcal{F}_T$- measurable that

$$\mathcal{E}(S)_T C^z = \mathbf{E}[\mathcal{E}(S)_T C^z] + \int_0^T \mathbf{E}[\mathcal{D}_v \mathcal{E}(S)_T C^z | \mathcal{F}_v] d\omega(v) \tag{1.159}$$

From Itô's rule it follows that

$$\mathcal{E}(-\int_0^{\cdot} (\alpha_s^z)^* dW_s^z)_T \mathcal{E}(S)_T C^z = \mathbf{E}[\mathcal{E}(S)_T C^z] + \int_0^T \mathcal{D}_v[\mathcal{E}(-\int_0^{\cdot} (\alpha_s^z)^* dW_s^z)_v \mathbf{E}[\mathcal{E}(S)_T C^z | \mathcal{F}_v]] dW_v^z \tag{1.160}$$

where we have used the commutativity of the conditional expectation and Malliavin derivative operator to get that

$$\mathcal{D}_v[\mathcal{E}(-\int_0^{\cdot} (\alpha_s^z)^* dW_s^z)_v \mathbf{E}[\mathcal{E}(S)_T C^z | \mathcal{F}_v]] = \mathcal{E}(-\int_0^{\cdot} (\alpha_s^z)^* dW_s^z)_v \mathbf{E}[\mathcal{D}_v[\mathcal{E}(S)_T C^z]| \mathcal{F}_v] - \mathcal{E}(-\int_0^{\cdot} (\alpha_s^z)^* dW_s^z)_v \mathbf{E}[\mathcal{E}(S)_T C^z | \mathcal{F}_v] (\alpha_v^z)^*$$

At the same time using again Itô's formula we have that

$$X_0^{\pi^z,c^z} + \int_0^T \mathcal{E}(-\int_0^{\cdot} (\alpha_s^z)^* dW_s^z)_v \mathcal{E}(S)_v ((\pi^z)_v^* \sigma_v - X_v^{\pi^z,c^z}(\theta_v + \alpha_v^z)^*) dW_v^z = \mathcal{E}(-\int_0^{\cdot} (\alpha_s^z)^* dW_s^z)_T \mathcal{E}(S)_T X_T^{\pi^z,c^z} + \int_0^T \mathcal{E}(-\int_0^{\cdot} (\alpha_s^z)^* dW_s^z)_v \mathcal{E}(S)_v (c_v^z - e_v) dv \tag{1.161}$$

From (1.160) and (1.161) we have for $(\hat{\pi}^z, \hat{c}^z)$ such that $\hat{c}^z_v = e_v$ and

$$(\hat{\pi}^z_v)^*\sigma_v - X^{\hat{\pi}^z,\hat{c}^z}_v(\theta_v + \alpha^z_v)^* = \frac{\mathcal{D}_v[\mathcal{E}(-\int_0^\cdot(\alpha^z_s)^* dW^z_s)_v \mathbf{E}[\mathcal{E}(S)_T C^z|\mathcal{F}_v]]}{\mathcal{E}(-\int_0^\cdot(\alpha^z_s)^* dW^z_s)_v \mathcal{E}(S)_v} \quad (1.162)$$

and initial wealth

$$X^{\hat{\pi}^z,\hat{c}^z}_0 = \mathbf{E}[\mathcal{E}(S)_T C^z] \quad (1.163)$$

that $\phi^{\hat{\pi}^z,\hat{c}^z,z}_T = 0$ for all $z \in \mathbb{R}^q$ $\mathbf{P}$- a.s.. Furthermore it follows that

$$X^{\hat{\pi}^z,\hat{c}^z}_t = \mathbf{E}^{\mathbf{P}^z}[\frac{\mathcal{E}(-\int_0^\cdot(\alpha^z_s)^* dW^z)_T \mathcal{E}(S)_T}{\mathcal{E}(-\int_0^\cdot(\alpha^z_s)^* dW^z)_t \mathcal{E}(S)_t} C^z|\mathcal{F}_t] \quad (1.164)$$

or since $\frac{d\mathbf{P}^z}{d\mathbf{P}}_{|\mathcal{F}_\cdot} = \mathcal{E}(\int_0^\cdot(\alpha^z_s)^* d\omega(s))$. from Bayes' law equivalently that

$$X^{\hat{\pi}^z,\hat{c}^z}_t = \mathbf{E}[\frac{\mathcal{E}(S)_T}{\mathcal{E}(S)_t} C^z|\mathcal{F}_t] \quad (1.165)$$

Using

$$\frac{\mathcal{D}_v\mathcal{E}(S)_T}{\mathcal{E}(S)_v} = \mathcal{D}_v\left[\frac{\mathcal{E}(S)_T}{\mathcal{E}(S)_t}\right] - \frac{\mathcal{E}(S)_T}{\mathcal{E}(S)_v}\theta^*_v \quad (1.166)$$

and therefore

$$\begin{aligned}\frac{\mathcal{D}_v[\mathcal{E}(-\int_0^\cdot(\alpha^z_s)^* dW^z_s)_v \mathbf{E}[\mathcal{E}(S)_T B^y|\mathcal{F}_v]]}{\mathcal{E}(-\int_0^\cdot(\alpha^z_s)^* dW^z_s)_v \mathcal{E}(S)_v} &= \mathbf{E}[\mathcal{D}_v[\frac{\mathcal{E}(S)_T}{\mathcal{E}(S)_v} B^z]|\mathcal{F}_v] \\ &\quad - \mathbf{E}[\frac{\mathcal{E}(S)_T}{\mathcal{E}(S)_v} B^z|\mathcal{F}_v](\theta_v + \alpha^z_v)^*\end{aligned}$$

we get from (1.162) that

$$(\hat{\pi}^z_t)^*\sigma_t = \mathbf{E}[\mathcal{D}_t \frac{\mathcal{E}(S)_T}{\mathcal{E}(S)_t} C^z|\mathcal{F}_t] \quad (1.167)$$

This establishes the results announced for an investor gambling upon the event $\{G = z\}$. If we evaluate (1.160) and (1.161) at $z = G$ we get for $(\hat{\pi}, \hat{c}) = (\hat{\pi}^G, \hat{c}^G)$ and $X^{\hat{\pi},\hat{c}}_0 = X^{\hat{\pi}^G,\hat{c}^G}_0$ that $\phi^{\hat{\pi},\hat{c},G}_T = 0$ $\mathbf{P}$- a.s. and from (1.167) we get (1.31). Finally if we apply lemma 23 to (1.165) we get since there is no tracking error the value of the claim given by (1.28)

Finally if $T \geq T_G$ and the signal reveals an event such that $B \neq \emptyset$ but $\mathbf{P}_G(B) = 0$ there is an arbitrage opportunity such as we have seen in theorem 8 that gains from trade at T_G are unbounded with probability one. Such an insider can therefore replicate any contingent claim at zero cost.

Q.E.D.

Proof of theorem 11

Proof

To establish the equivalence of implicit prices of insiders and outsiders it is sufficient to show that $\tilde{\mathbf{Q}} = \mathbf{Q}$ on $\mathcal{F}_T$. Since for all $E \in \mathcal{F}_T$

$$\tilde{\mathbf{Q}}(E) = \mathbf{E}^{\mathbf{P}_G}[\mathbf{E}^{\mathbf{P}^z}[\mathbf{1}_E \frac{d\mathbf{P}}{d\mathbf{P}^z}_{|\mathcal{F}_T} \frac{d\mathbf{Q}}{\mathbf{P}}_{|\mathcal{F}_T}]] \tag{1.168}$$

and

$$\mathbf{E}^{\mathbf{P}^z}[\mathbf{1}_E \frac{d\mathbf{P}}{d\mathbf{P}^z}_{|\mathcal{F}_T} \frac{d\mathbf{Q}}{\mathbf{P}}_{|\mathcal{F}_T}] = \mathbf{E}[\mathbf{1}_E \frac{d\mathbf{Q}}{\mathbf{P}}_{|\mathcal{F}_T}] \tag{1.169}$$

we have that

$$\tilde{\mathbf{Q}}(E) = \mathbf{E}^{\mathbf{P}_G}[\mathbf{Q}(E)] \tag{1.170}$$

and therefore $\tilde{\mathbf{Q}}(E) = \mathbf{Q}(E)$.

Q.E.D.

1.10.3 Proofs of section 1.5

Proof of theorem 12

First it follows from theorem 8 that if insiders have no arbitrage opportunities investment horizons must end before resolution $T < T_G$ and/or insider information

$\sigma(G)$ is completely atomic and consequently $\tilde{\mathbf{Q}} \ll \mathbf{P}$ on $\mathcal{G}_t$ for all $t \in [0,1]$ and therefore also $\mathbf{P}^z \ll \mathbf{P}$ on $\mathcal{F}_t$ for all $t \in [0,1]$. It follows that in the absence of arbitrage we have

$$\mathbf{P}^z(\{\frac{q(z)}{p(\omega,t,z)} \in]0,+\infty[\}) = 1 \tag{1.171}$$

Then if we define $e^z := e_0 + \mathbf{E}^{\mathbf{P}^z}[\int_0^T \mathcal{E}(S^z)_t e_t dt]$ and $\mathcal{E}(S^z) := \mathcal{E}(S)_t \frac{q(z)}{p(\omega,t,z)}$ we have for fixed $z \in \mathbb{R}^q$ that the value function

$$J(e^z) := \inf_{y^z>0} \mathbf{E}^{\mathbf{P}^z}[\int_0^T u(t, I(t, y^z\mathcal{E}(S^z)_t)) - y^z\mathcal{E}(S^z)_t I(t, y^z\mathcal{E}(S^z)_t)dt + y^z e^z] \tag{1.172}$$

satisfies

$$J(e^z) = \mathbf{E}^{\mathbf{P}^z}[\int_0^T u(t, I(t, \hat{y}^z\mathcal{E}(S^z)_t))dt] \tag{1.173}$$

where $\hat{y}^z$ is such that

$$\mathbf{E}^{\mathbf{P}^z}[\int_0^T \mathcal{E}(S^z)_t I(t, \hat{y}^z\mathcal{E}(S^z)_t))dt] = e^z \tag{1.174}$$

and where the existence of $\hat{y}^z$ follows from (1.35). Now since for $\mathcal{F}_t$ adapted non-negative consumption processes c

$$\mathbf{E}^{\mathbf{P}^z}[\int_0^T \sup_{c\geq 0}[u(t,c_t) - \hat{y}^z c_t]dt] \geq \sup_{c^z\geq 0} \mathbf{E}^{\mathbf{P}^z}[\int_0^T [u(t,c_t) - \hat{y}^z c_t]dt] \tag{1.175}$$

and the convex conjugate function in the first integral is

$$\sup_{c^z\geq 0}[u(t,c_t) - \hat{y}^z\mathcal{E}(S^z)_t c_t] = u(t, \hat{c}_t^z) \tag{1.176}$$

we have established that the consumption policy of a "gambler"

$$\hat{c}_t^z = I(t, \hat{y}^z\mathcal{E}(S^z)_t) \tag{1.177}$$

is optimal for the problem

$$\sup_{c\geq 0} \mathbf{E}^{\mathbf{P}^z}[\int_0^T u(t,c_t)dt] \tag{1.178}$$

subject to the static budget constraint

$$\mathbf{E}^{\mathbf{P}^z}[\int_0^T \mathcal{E}(S^z)_t c_t dt] = e^z \tag{1.179}$$

where $\hat{y}^z$ is the multiplier associated to the constraint. To establish that these strategies are optimal for the dynamic problem

$$\sup_{(\pi,c)\in\mathcal{A}(\mathbf{P}^z,\mathbb{F},e)} \mathbf{E}^{\mathbf{P}^z}[\int_0^T u(t,c_t)dt] \tag{1.180}$$

it is sufficient to show that optimal wealth satisfies $X^{\hat{\pi}^z,\hat{c}^z} > -K$ some $K > 0$ and $X_T^{\hat{\pi}^z,\hat{c}^z} \geq 0$ both $\mathbf{P}^z$- a.s.. Since preferences are strictly monotone and there is no incentive for a bequest we have that $X_T^{\hat{\pi}^z,\hat{c}^z} = 0$ $\mathbf{P}^z$-a.s.. Then since

$$\mathbf{E}^{\mathbf{P}^z}[\mathcal{E}(S^z)_T X_T^{\pi,c}|\mathcal{F}_t] - \mathcal{E}(S^z)_t X_t^{\pi,c} = \mathbf{E}^{\mathbf{P}^z}[\int_0^T \mathcal{E}(S^z)_t(c_t - e_t)dt|\mathcal{F}_t] \tag{1.181}$$

it follows from the fact that the endowment process e is bounded from below that the optimal wealth $X^{\hat{\pi}^z,\hat{c}^z}$ must be bounded from below. This establishes the optimality of $\hat{c}^z$ for an investor "gambling" upon the event $\{G = z\}$.

Next we have to show that $\hat{c}^G$ is optimal for an investor having beliefs $\mathbf{P}$ and additional information $\sigma(G)$.

Since all $c_t \in \mathcal{A}(\mathbf{P},\mathbb{G},e)$ are also in $\mathbb{L}^{1,2}(\mathbb{R}^d)$ we have from theorem 22 that there exists $c^z = (c_t^z; t \in [0,T])$ such that

$$c_t = \tilde{c}_t^G \tag{1.182}$$

and where $c^z \in \mathcal{A}(\mathbf{P}^z,\mathbb{F},e)$. From the optimality of $\hat{c}^z$ in $\mathcal{A}(\mathbf{P}^z,\mathbb{F},e)$ we must have that

$$\mathbf{E}^{\mathbf{P}^z}[\int_0^T u(t,\hat{c}_t^z)dt] \geq \mathbf{E}^{\mathbf{P}^z}[\int_0^T u(t,c_t^z)dt] \tag{1.183}$$

for fixed $z \in \mathbb{R}^q$. This remains true at $z = G$ and it follows from the lemma 23 that

$$\mathbf{E}[\int_0^T u(v,\hat{c}_v^G)dv|\mathcal{G}_0] \geq \mathbf{E}[\int_0^T u(v,c_v)dv|\mathcal{G}_0] \tag{1.184}$$

for all $c_t \in \mathcal{A}(\mathbf{P}, \mathbb{G}, e)$.

It remains to proof that the cumulative net consumption process $\int_0^T(\hat{c}_t - e_t)dt$ can be financed by $\mathcal{G}_t$- measurable portfolio strategy. For $T < T_G$ this was established in theorem 10. We have to show that the same kind of argument remains true for discrete signals if $T \geq T_G$. We have already established that $\frac{q}{p(\omega,t,z)} < \infty$ for all $t \in [0,1]$ $\mathbf{P}^z$- a.s.. In fact we have on $\{G = z\}$ for $t \in [\![T_G, 1]\!]$

$$\mathcal{E}(S^z)_t = \mathbf{P}_G(\{z\})\mathcal{E}(S)_t \tag{1.185}$$

$\mathbf{P}^z$-a.s. and therefore that the density process of the absolutely continuous local martingale measure $\tilde{\mathbf{Q}}$ is finite and positive on $\{G = z\}$. We therefore have on $\{G = z\}$ from the comparison of

$$\int_0^T \mathcal{E}(s^z)_t(\hat{c}_t^z - e_t)dt = e_0 + \int_0^T \mathcal{E}(S^z)_t(\pi_t^*\sigma_t - (\theta_t + \alpha_t^z)^*)dW_t^z \tag{1.186}$$

where $\alpha_t^z = 0$ and correspondingly $dW_t^z = d\omega(t)$ for $t \in [\![T_G, 1]\!]$ $\mathbf{P}^z$-a.s. with

$$\int_0^T \mathcal{E}(S^z)_t(\hat{c}_t^z - e_t)dt = e_0 + \int_0^T \mathcal{D}_t\left\{\frac{q(z)}{p(\omega,t,z)}\mathbf{E}[\int_t^T \mathcal{E}(S)_t(\hat{c}_t^z - e_t)dt|\mathcal{F}_t]\right\}dW_t^z$$

that

$$\mathcal{D}_t\left\{\frac{q(z)}{p(\omega,t,z)}\mathbf{E}[\int_0^T \mathcal{E}(S)_t(\hat{c}_t^z - e_t)dt|\mathcal{F}_t]\right\} = \mathcal{E}(S^z)_t(\hat{\pi}_t^z)^*\sigma_t - (\theta_t + \alpha_t^z)^*) \tag{1.187}$$

Since $\mathbf{P}^z(\{G = z\}) = 1$ solving for $\hat{\pi}_t^z$ as in the proof of theorem 10 gives the results on $\mathbf{P}^z \otimes \lambda$-a.e.. As for consumption expressions for optimal strategies of a gambler "evaluated" at $z = G$ are optimal for the consumption-investment problem of an insider since they replicate $\mathcal{G}_t$ adapted cumulative net-consumption without tracking error.

Finally if there are arbitrage opportunities then as we have seen in theorem 8 we can finance any cumulative consumption process with zero cost. This implies that the static budget constraint will never bind and therefore the marginal value of wealth $\hat{y}^z$ is zero. It follows that the first order condition $\partial_2 u(t, \hat{c}_t^z) = 0$. But this is by the Inada conditions only possible if consumption is unbounded with positive $\mathbf{P}^z$- probability. Since $\mathbf{P}^z \not\ll \mathbf{P}$ such states have non-zero $\mathbf{P}^z$- probability and consequently the value of the problem for fixed outcome of the signal $G = z$ explodes. This remains true if we "evaluate" at $z = G$.

Q.E.D.

Proof of proposition 13 and corollaries 14, 15 and 16

Again we are first conditioning on $\{G = z\}$ in a first step and get results by "evaluating" at $z = G$ in a second step. Since the marginal rate of substitution of a gambler can be written for the state dependent Bernoulli indicator $v(\omega, \cdot, \cdot)$ in (1.53) as

$$\frac{\partial_2 u(s, \hat{c}_s^z)}{\partial_2 u(t, \hat{c}_t^z)} = \left(\frac{\left(\frac{d\mathbf{P}}{d\mathbf{P}^z}\right)_t}{\left(\frac{d\mathbf{P}}{d\mathbf{P}^z}\right)_s} \right)(\omega) \frac{\partial_2 v(\omega, s, \hat{c}_s^z)}{\partial_2 v(\omega, s, \hat{c}_s^z)} \tag{1.188}$$

we can on $\{G = z\}$ equivalently write (1.47) as

$$\begin{aligned}(\hat{\pi}_t^z)^* \sigma_t = \mathbf{E}[\int_t^T \frac{\partial_2 v(\cdot, u, \hat{c}_u^z)}{\partial_2 v(\cdot, t, \hat{c}_t^z)}(\hat{c}_u^z - e_u) du | \mathcal{F}_t] A(t, \hat{c}_t^z) \mathcal{D}_t \hat{c}_t^z + \\ \mathbf{E}[\int_t^T \frac{\partial_2 v(u, \hat{c}_u^z, \cdot, z)}{\partial_2 v(t, \hat{c}_t^z, \cdot, z)}[1 - (\hat{c}_u^z - e_u) A(u, \hat{c}_u^z)] \mathcal{D}_t \hat{c}_u du | \mathcal{F}_t] - \\ \mathbf{E}[\int_t^T \frac{\partial_2 v(u, \hat{c}_u^z, \cdot, z)}{\partial_2 v(t, \hat{c}_t^z, \cdot, z)} \mathcal{D}_t e_u) du | \mathcal{F}_t]\end{aligned} \tag{1.189}$$

But from the expression for "gambler's" optimal consumption policy (1.177) taking instantaneous Malliavin derivatives we get

$$\mathcal{D}_t \hat{c}_t^z = \frac{1}{A(t, \hat{c}_t^z)}(\alpha^z + \theta)^* \tag{1.190}$$

and for $t < u$

$$\mathcal{D}_t \hat{c}_u^z = \frac{1}{A(u, \hat{c}_u^z)} \mathcal{D}_t [\log \frac{q(z) \mathcal{E}(S)_u}{p(\omega, t, z)}] \tag{1.191}$$

Finally it follows from theorem 2.1.1 page 102 of Nualart (1995) that the Malliavin derivative of the endowment rate process $\mathcal{D}_t e_u$ is found as solution of the linearized stochastic differential equation(1.59). If we replace this expression together with (1.190) and (1.191) in (1.189) we get the result announced by arranging terms.

Q.E.D.

Proof of proposition 17

In the proofs of corollaries 2 we have shown that (1.68) implies $\alpha_t^z = 0$ $\mathbf{P} \otimes \boldsymbol{\lambda}$- a.e.. It follows from the expressions of optimal consumption in theorem 12 that in this case $\hat{c}_t^G = \hat{c}_t^z = \tilde{c}_t$ for all $t \in [0, T]$ where $\tilde{c}$ denotes the optimal consumption policy in $\mathcal{A}(\mathbf{P}, \mathbb{F}, e)$. This is sufficient to establish the equivalence of the value functions.

Q.E.D.

Proof of proposition 18

Since admissible strategies in the consumption-investment problem of an insider have to be $\mathcal{G}_t$- adapted we clearly have that

$$\mathcal{F}_t^{\hat{c}^G} \bigvee \mathcal{F}_t^{\hat{\pi}^G} \subset \mathcal{G}_t \tag{1.192}$$

for all $t \in [0, T]$ and it remains to find conditions for which we have the reversed inclusion. From (1.43) we see that $\hat{\pi}_0^G = 0$ and therefore necessary conditions for which the information generated by optimal strategies corresponds to all individual private information must also guarantee that

$$\sigma(\hat{c}_0^G) = \sigma(G) \tag{1.193}$$

But to establish this is equivalent to establish the existence of a Borel function such that

$$G = h(\hat{c}_0^G) \tag{1.194}$$

Since $\hat{c}_0^G$ satisfies the state by state first order conditions

$$\partial_2 u(0, \hat{c}_0^G) = \mathcal{Y}(e^G, G) \tag{1.195}$$

where $\mathcal{Y}(x,z)$ is such that $\mathcal{X}(\mathcal{Y}(x,z),z) = x$ it is by the definition of $\mathcal{X}(y,z)$ sufficient for the existence of the Borel measurable function h to show that the mapping

$$z \mapsto \mathbf{E}[\int_0^T \mathcal{E}(S)_u e_u \frac{q(z)}{p(\omega,u,z)})du] \tag{1.196}$$

is bijective. It follows that for h to exist the mapping

$$z \mapsto \frac{q(z)}{p(\omega,u,z)} \tag{1.197}$$

must be bijective $\mathbf{P} \otimes \boldsymbol{\lambda}$- a.e.. Then suppose such a mapping exists and is given by k. Consequently it must be true that $k^{-1}(z)p(\omega,t,z) = q(z)$ $\mathbf{P} \otimes \boldsymbol{\lambda}$- a.e.. This would imply since $\mathbf{E}[p(\omega,t,z)] = q(z)$ that $k^{-1}(z) = 1$ which proofs that such a bijection does not exist unless the signal is independent of the public information. Consequently the optimal insider strategies will never be fully revealing if it is non-redundant.

Q.E.D.

Chapter 2

A Monte Carlo Method for Optimal Portfolios

2.1 Introduction

Asset allocation models have received extensive attention in the past three decades. Prompted by the seminal work of Merton (1969, 1971) researchers have explored various aspects of the problem in the context of financial markets with diffusion price processes (e.g. Richard (1975)). Numerical methods based on the dynamic programming approach employed in this literature have also been used to examine the properties of optimal portfolios (Brennan, Schwarz and Lagnado (1997)). Numerical schemes based on PDEs, however, become increasingly difficult to implement when the number of underlying state variables increases. More recent contributions by Karatzas, Lehoczky and Shreve (1987) and Cox and Huang (1989) have proposed an alternative resolution method based on martingale techniques. This approach yields a closed form solution for optimal consumption when markets are complete even when asset prices follow Ito processes with history-dependent coefficients. The optimal portfolio was derived by Ocone and Karatzas (1991) using the Clark-Ocone formula. This representation formula expresses the portfolio in terms of expectations of random variables which involve "abstract" Malliavin derivatives of the coefficients of the model, namely the interest rate (IR) and the market price of risk (MPR).

But while theoretical formulas for optimal portfolios are available in general contexts little is known about the structure and properties of the hedging components. Even if we restrict attention to diffusion models, realistic specifications with stochastic IR and MPR give rise to complex hedging terms which depend on multiple state variables and are often difficult to evaluate numerically. As a result attention has been devoted to (i) state variable specifications for which closed form solutions are available (Kim and Omberg (1996), Liu (1999), Wachter (1999)) or (ii) specifications which are computationally tractable based on dynamic programming techniques (Brennan, Schwarz and Lagnado (1997)), or (iii) discrete time models based on approximated Euler equations (Campbell and Viceira (1999)).

This paper provides three main contributions. First we exploit the diffusion nature of the opportunity set to provide explicit expressions for the Malliavin derivatives arising in the hedging components of the optimal portfolio. Hedging demands are expressed as conditional expectations of random variables which depend on the drift and variance of the relevant state variables. These formulas are valid for any structure of the underlying processes and of the utility function and reduce the computation of hedging demands to the computation of expectations, as in traditional option pricing. Our approach can therefore be seen as a translation of the dynamic asset allocation problem into an option pricing problem for which Monte Carlo methods, as summarized in Boyle, Broadie and Glasserman (1997), have long been successfully applied by practitioners[1]. Furthermore, the formulas enable us to establish new theoretical results about the hedging behavior.

[1]See also Fournié et. al. (1999)

Second we derive an alternative representation of Malliavin derivatives of diffusion processes which simplifies their evaluation. Our formula relies on a variance-stabilizing transformation of the underlying process and eliminates stochastic integrals from their representation. Aside from its theoretical interest this new expression has interesting computational benefits. Indeed, the absence of stochastic integrals ensures the existence of an exact weak approximation scheme for the martingale part of the Malliavin derivatives and this improves the rate of convergence of approximations of Malliavin derivatives to their true values. The scheme also increases the speed of convergence of simulated trajectories of hedging terms and of any statistic (such as confidence intervals) of simulated hedging terms. Finally it may also help to reduce the second-order bias and therefore the size distortion of asymptotic confidence intervals of the Monte Carlo estimator of the hedging demands and portfolios given the realization of the state variables.

Third we provide new results on the economic properties of optimal portfolios. We examine bivariate and trivariate IR and MPR models in a setting with constant relative risk aversion. In our benchmark bivariate model the IR process is mean-reverting with square-root volatility (MRSR) and the MPR process is Gaussian with either mean-reversion (MRG) or with mean-reversion and interest rate dependence in the drift (MRGID). More elaborate trivariate models with stochastic dividend yield or volatility, and with multiple assets are also considered. In these settings we document the magnitude of the hedging terms and their behavior relative to the parameters of the model such as risk aversion, investment horizon or IR and MPR values. All our results are based on a portfolio formula which evolves from the Ocone-Karatzas representation. This modified formula which emphasizes the role of relative risk aversion and wealth sheds further light on the portfolio/hedging behavior. It can be viewed as a minor contribution of the paper.

Some of the lessons drawn from our simulations can be summarized as follows:

1. Our methodology involving the combination of Monte-Carlo simulation and our variance-stabilizing transformation produces very reasonable values for the shares of wealth invested in the stock. Unlike some earlier studies of optimal portfolios interior solutions are obtained and portfolio shares are stable in simulation exercises such as market timing experiments.

2. Hedging components are important for asset allocation purposes. For long horizons the adjustment to mean-variance demands can represent up to 80% of the stock demand. Hedging demands also exhibit low volatility and are therefore very stable over time.

3. Critical factors in optimal asset allocation are the risk aversion and the investment horizon of the investor. For instance, in our basic bivariate model, investors with short (long) horizons and whose risk aversion exceeds 1 want to reduce (increase) their stock demand relative to the logarithmic investor in order to hedge against MPR (IR) fluctuations. The effects documented in the paper rationalize the marketing of investment products tailored to different categories of investors classified according to those criteria.

4. Allocation rules are remarkably stable relative to the other parameters of the model. Variations of the order of 2 standard deviations around estimated parameter values have little impact on the magnitude of investment shares.

5. The global behavior of the optimal portfolio in the multiasset case parallels the behavior displayed with a single risky asset. Hedging terms exhibit strong patterns with respect to correlations when asset returns are highly correlated. Correlations between returns and among state variables emerge as additional factors driving the size of hedging demands.

The portfolio choice problem is stated next. Section 3 presents a closed-form solution and discusses its structure. Section 4 develops an alternative formula for Malliavin derivatives of diffusion processes. Numerical implementation is discussed in section 5. Our basic bivariate model with MRSR interest rate and MRG/MRGID market price of risk is analyzed in sections 6 and 7. Sections 8 and 9 provide trivariate extensions to stochastic dividends and stochastic, imperfectly correlated volatility. A multiasset model is analyzed in section 10. Proofs are in appendix A; appendix B extends the procedure to multivariate diffusions; appendix C contains results for the MRGID model; appendix D reports asymptotic properties of state variable estimators.

2.2 The portfolio choice problem

We consider a portfolio choice problem in an economy with d state variables $Y_{jt}, j = 1, ..., d$, and d sources of Brownian uncertainty $W_{it}, i = 1, ..., d$.[2] State variables follow the vector diffusion process

$$dY_t = \mu^Y(t, Y_t)dt + \sigma^Y(t, Y_t)dW_t \tag{2.1}$$

where the coefficients satisfy appropriate Growth and Lipschitz conditions for the existence of a unique strong solution.[3] The investor allocates his wealth between d risky securities and one riskless asset (a money market account) with instantaneous riskless rate of return $r_t = r(t, Y_t)$. The security prices $S_i, i = 1, ..., d$, satisfy the stochastic differential equations

$$dS_{it} = S_{it}[(\mu_i(t, Y_t) - \delta_i(t, Y_t))dt + \sigma_i(t, Y_t)dW_t]; \quad 1 \leq i \leq d \tag{2.2}$$

[2]It is straightforward to consider $k \neq d$ state variables. To simplify notation, in particular for the expressions of the Malliavin derivatives, we assume that $k = d$.

[3]Note that the d state variables are joint solutions of the system (3.2), i.e. they influence each other. Remark 1 considers the special case of an autonomous system in which each state variable is determined independently.

where μ_i is the expected return, δ_i the dividend rate and σ_i the vector of volatility coefficients of security i. We assume that $r(t, Y_t), \mu_i(t, Y_t), \delta_i(t, Y_t)$ are integrable $(P - a.s.)$ and that $\sigma_i(t, Y_t)$ is square-integrable $(P - a.s.)$. Let σ denote the $d \times d$-dimensional volatility matrix whose rows are $\sigma_i, i = 1, ..., d$. Suppose that σ is nonsingular almost everywhere and that the market price of risk

$$\theta_t = \theta(t, Y_t) = \sigma(t, Y_t)^{-1}(\mu(t, Y_t) - r(t, Y_t)\mathbf{1}),$$

where $\mathbf{1}$ is the unit vector, is continuously differentiable and satisfies the Novikov condition $E \exp\left(\frac{1}{2}\int_0^T \theta_t'\theta_t dt\right) < \infty$. Under this condition the risk neutral measure is well defined and given by $dQ = \eta_T dP$ where

$$\eta_t = \exp\left[-\int_0^t \theta_t' dW_t - \frac{1}{2}\int_0^t \theta_t'\theta_t dt\right].$$

The state price density is $\xi_t \equiv B_t^{-1}\eta_t$ where $B_t \equiv \exp[\int_0^t r_s ds]$ is the date t-value of a dollar investment in the money market account. Relative state prices are written $\xi_{t,v} \equiv \xi_v/\xi_t$. Under Q the process $W_t^{\mathbf{Q}} = W_t + \int_0^t \theta_v dv$ is a Brownian motion.

Suppose that an investor seeks to maximize the expected utility of his terminal wealth by selecting a dynamic portfolio policy composed of the d risky assets and the riskless asset

$$\max_{\pi} U(X_T) \equiv E[u(T, X_T)] \quad s.t. \tag{2.3}$$

$$\begin{cases} dX_t = r_t X_t dt + \pi_t'[(\mu_t - r_t\mathbf{1})dt + \sigma_t dW_t], & X_0 = x \\ X_t \geq 0 \text{ for all } t \in [0, T]. \end{cases} \tag{2.4}$$

Here X_t represents the investor's wealth at date t, x is his initial wealth and π_t the amounts invested in the risky assets at date t. The nonnegativity constraint is a typical no-bankruptcy condition. The utility function is strictly increasing and concave with limiting values $\lim_{X\to\infty}\partial_2 u(T,x)=0$ and $\lim_{X\to 0}\partial_2 u(T,x)=\infty$ for all $T<\infty$. (For any function $f(t,X)$ we write $\partial_i f$ for the first derivative relative to i, $i=1,2$ and $\partial_{ij}f$ the second derivative, $i,j=1,2$; when the second argument is a vector $\partial_2 f$ is the gradient and $\partial_{22}f$ the hessian of second derivatives).

2.3 The optimal portfolio: the investor's hedging behavior

The portfolio choice problem described above can be resolved by using a martingale approach (Karatzas, Lehoczky and Shreve (1987), Cox-Huang (1989)) to identify optimal terminal wealth in explicit form and then applying the Clark-Ocone formula on the representation of Brownian functionals to obtain the financing portfolio. This approach was adopted by Ocone and Karatzas (1991) who provide formulas in the form of conditional expectations of random variables involving Malliavin derivatives. Due to the generality of their model in which asset prices follow Ito processes (with unspecified coefficients) these Malliavin derivatives are abstract quantities without an explicit structure. In this section we exploit the diffusion specification of the financial market to derive explicit expressions for the Malliavin derivatives and hence for the optimal portfolio.

2.3.1 The optimal portfolio policy.

Let $V(x)$ denote the value function in the optimization problem (3.3)-(3.4), $I(T,y)$ the inverse marginal utility, $\widehat{y}$ the marginal value of initial wealth and $\hat{X}$ the optimal wealth. Our first result identifies the general structure of the optimal portfolio and of its hedging components.

Theorem 24 *If $V(x) < \infty$ and $\xi_T I(T, \widehat{y}\xi_T) \in \mathbb{D}^{1,2}$ we have that*[4]

$$\begin{aligned} \hat{\pi}_t &= \hat{X}_t \frac{1}{R(t,\hat{X}_t)}(\sigma(t,Y_t)')^{-1}\theta(t,Y_t)c(t,Y_t) \\ &\quad + \hat{X}_t(\frac{1}{R(t,\hat{X}_t)} - 1)(\sigma(t,Y_t)')^{-1}a(t,Y_t) \\ &\quad + \hat{X}_t(\frac{1}{R(t,\hat{X}_t)} - 1)(\sigma(t,Y_t)')^{-1}b(t,Y_t) \end{aligned} \tag{2.5}$$

where $R(t,x) := \frac{-\partial_{22}u(t,x)x}{\partial_2 u(t,x)}$ denotes the Arrow-Pratt measure of relative risk aversion, and

$$a(t,Y_t)' \equiv \mathbf{E}_t^{\mathbf{Q}}\left[\frac{\frac{\hat{X}_T}{B_T}}{\frac{\hat{X}_t}{B_t}}\left(\frac{1-1/R(T,\hat{X}_T)}{1-1/R(t,\hat{X}_t)}\right)\int_t^T \mathcal{D}_t r_s ds\right] \tag{2.6}$$

$$b(t,Y_t)' \equiv \mathbf{E}_t^{\mathbf{Q}}\left[\frac{\frac{\hat{X}_T}{B_T}}{\frac{\hat{X}_t}{B_t}}\left(\frac{1-1/R(T,\hat{X}_T)}{1-1/R(t,\hat{X}_t)}\right)\int_t^T (dW_s^{\mathbf{Q}})'\mathcal{D}_t\theta_s\right] \tag{2.7}$$

$$c(t,Y_t) \equiv \mathbf{E}_t^{\mathbf{Q}}\left[\frac{\frac{\hat{X}_T}{B_T}}{\frac{\hat{X}_t}{B_t}}\frac{R(t,\hat{X}_t)}{R(T,\hat{X}_T)}\right]. \tag{2.8}$$

In these expressions optimal wealth equals $\hat{X}_t = \mathbf{E}_t[\xi_{t,T}I(T,\widehat{y}\xi_T)]$. The Malliavin derivatives in (3.6)-(3.7) are given in explicit form by $\mathcal{D}_t\theta_s' = \partial_2\theta(s,Y_s)'\mathcal{D}_tY_s$ and $\mathcal{D}_t r_s = \partial_2 r(s,Y_s)\mathcal{D}_tY_s$ where

$$\mathcal{D}_tY_s = \sigma^Y(t,Y_t)\exp\left\{\int_t^s dL_v\right\}, \tag{2.9}$$

[4] $\mathbb{D}^{1,2}$ is the domain of the Malliavin derivative. See Nualart (1995) for exact definitions.

with the $d \times d$ random variable dL_v defined by[5]

$$dL_v \equiv \left[\partial_2 \mu^Y(v, Y_v) - \frac{1}{2} \sum_{j=1}^{d} \partial_2 \sigma_{\cdot j}^Y(v, Y_v)(\partial_2 \sigma_{\cdot j}^Y(v, Y_v))' \right] dv + \sum_{j=1}^{d} \partial_2 \sigma_{\cdot j}^Y(v, Y_v) dW_{jv} \tag{2.10}$$

where $\sigma_{\cdot j}^Y$ denotes the j^{th} column of the matrix σ^Y.

Note that the first component of the optimal portfolio (2.5) is a mean-variance component while the two other components are intertemporal hedging terms (see Merton (1971)).[6] In this general formula the mean-variance term varies with optimal wealth since the coefficient of relative risk aversion is allowed to change with wealth. Hedging arises since the investor seeks insurance against fluctuations in the interest rate (second component of (2.5)) and in the market prices of risk (third component of (2.5)). That the second term is motivated by the desire to hedge interest rate risk is evidenced by the presence of the Malliavin derivative $\mathcal{D}_t r_s$ which captures the interest rate's sensitivity to the underlying risk factors, i.e. the Brownian motion processes W_i. In accordance we call this term an IR-hedge.[7] Similarly the third term is seen to emerge when the market prices of risk are sensitive to the W_i (i.e. when $\mathcal{D}_t \theta_s \neq 0$) and is called an MPR-hedge. When (r, θ) are constant or deterministic all these hedging terms are null since in the Malliavin derivatives $\mathcal{D}_t r_s$ and $\mathcal{D}_t \theta_s$, the partial derivatives $\partial_2 r(s, Y_s)$ and $\partial_2 \theta(s, Y_s)$ are zero.

[5]The exponential in (2.9) should be interpreted as the exponential of a matrix, i.e. (2.9) is short hand notation for the solution of $d\mathcal{D}_t Y_s = \left(dL_s + \frac{1}{2} d[L]_s\right) \mathcal{D}_t Y_s$ subject to the boundary condition $\mathcal{D}_t Y_t = \sigma^Y(t, Y_t)$, where $[L]$ is the quadratic variation process.

[6]The optimal portfolio formula extends easily to the case of intermediate consumption. It also extends to settings with infinite horizon provided that the Novikov condition is satisfied.

[7]Expression (2.9) shows that the Malliavin derivative, in a Markovian model, corresponds to the derivative of the stochastic flow of the SDE of state variables with respect to the initial position of the state variables (Colwell, Elliott and Kopp (1991)).

Before discussing the behavior embedded in the hedging components it is also of interest to point out that formula (2.5) expresses the hedging components in explicit form: hedging demands are conditional expectations of random variables which depend entirely on the exogenous coefficients of the model and the utility function. The key to these explicit expressions is the derivation of closed-form solutions for the Malliavin derivatives $\mathcal{D}_t r_s$ and $\mathcal{D}_t \theta_s$ which are obtained due the diffusion structure of the uncertainty. As mentioned above these results complement Ocone and Karatzas (1991) who express the optimal portfolio in terms of abstract Malliavin derivatives.

2.3.2 The intertemporal hedging behavior.

Let us now focus on the hedging behavior of the investor. First, it should be noted that a myopic individual ($R(t, \widehat{X}_t) = 1$) does not hedge[8]. The signs of the hedging terms will otherwise depend on the signs of the conditional expectations $a(t, Y_t)$ and $b(t, Y_t)$. For example, when these are positive, an individual who is more (less) risk tolerant than the logarithmic investor will over- (under) invest in the risky assets. For the IR-hedge simple sufficient conditions ensure an unambiguous behavior.

Proposition 25 *Fix $t \in [0, T]$. Suppose that the conditions*

(i) $(\sigma(t, Y_t)')^{-1}(\mathcal{D}_t r_s)' \leq 0$ for all $s \geq t$,(P-a.s)

(ii) $R(t, \hat{X}_t) \geq 1$ and $R(T, \hat{X}_T) \geq 1$ (P-a.s).

hold. Then, intertemporal hedging of interest rate risk raises the demand for stocks (i.e. the IR-hedge is nonnegative). If (i)-(ii) hold for all $t \in [0, T]$ the IR-hedge boosts the proportion of wealth invested in stocks at all times.

[8]When $R(t, \widehat{X}_t)$ and $R(T, \widehat{X}_T)$ tend to one, the ratio inside the conditional expectations (3.6)-(3.7) tends to one as seen by applying l'Hôpital's rule.

Conditions (i)-(ii) are very general. The first condition holds in a variety of special cases that are of interest for empirical or theoretical reasons. For instance it holds if state variables are autonomous (see remark 1 below) and

$$(\sigma(t,Y_t)')^{-1}\left(\partial_2 r(t,Y_t)\sigma^Y(t,Y_t)\right)' \leq 0. \tag{2.11}$$

In the single risky asset case this simply boils down to negative correlation between the interest rate and the risky asset price, which is empirically verified if the risky asset is interpreted as the SP500 index. Condition (2.11) also holds with multiple risky assets that are independent and negatively correlated with the interest rate. In all these cases the particular structure of the coefficients of the state variables processes (whether they are increasing, decreasing, convex or concave functions) does not matter for the sign of the hedging term: the only aspect of relevance is whether (2.11) is verified.

The second condition applies even to models in which relative risk aversion varies with optimal wealth. As long as an investor displays more risk aversion than a myopic investor at date t and for all possible realizations of optimal terminal wealth the condition will hold.

When we combine both conditions we obtain, for instance, the intuitive proposition that individuals that are more risk averse than the log investor ($R(t,\hat{X}_t) \geq 1, R(T,\hat{X}_T) \geq 1$) will increase their demand for the market portfolio of risky assets when the interest rate covaries negatively with the portfolio return (single risky asset model) in order to hedge interest rate risk.

We conclude this discussion with a description of hedging demands and Malliavin derivatives for the case of autonomous state variables.

Remark 1 *When the system of stochastic differential equations (3.2) is composed of d autonomous equations $dY_{it} = \mu_i^Y(t,Y_{it})dt + \sigma_i^Y(t,Y_{it})dW_t$ for $i = 1,...,d$, we* *can write*

$$\mathcal{D}_t Y_{is} = \sigma_i^Y(t, Y_{it}) \exp\left\{\int_t^s dL_v^i\right\} \tag{2.12}$$

$$dL_v^i \equiv [\partial_2 \mu_i^Y(v, Y_{iv}) - \frac{1}{2}\partial_2 \sigma_i^Y(v, Y_{iv})(\partial_2 \sigma_i^Y(v, Y_{iv}))']dv + \partial_2 \sigma_i^Y(v, Y_{iv})dW_v \tag{2.13}$$

In this instance the sign of the Malliavin derivative (3.8) is positive (negative) when $\sigma_i^Y(t, Y_{it})$ is positive (negative). The results discussed above follow immediately from this property.

2.3.3 Constant relative risk aversion.

Since our numerical results in later sections assume constant relative risk aversion we specialize the formulas of theorem 34 to that case.

Theorem 26 *(CRRA) When the utility function exhibits constant relative risk aversion R the optimal portfolio is*

$$\hat{\pi}_t = \hat{X}_t(\sigma(t, Y_t)')^{-1}\left[\frac{1}{R}\theta(t, Y_t) + (\frac{1}{R} - 1)a(t, Y_t) + (\frac{1}{R} - 1)b(t, Y_t)\right] \tag{2.14}$$

where

$$a(t, Y_t)' \equiv \mathbf{E}_t^{\mathbf{Q}}\left[\frac{\frac{\hat{X}_T}{B_T}}{\frac{\hat{X}_t}{B_t}}\int_t^T \mathcal{D}_t r_s ds\right] = \frac{\mathbf{E}_t\left[\xi_{t,T}^{1-1/R}\int_t^T \mathcal{D}_t r_s ds\right]}{\mathbf{E}_t\left[\xi_{t,T}^{1-1/R}\right]} \tag{2.15}$$

$$b(t, Y_t)' \equiv \mathbf{E}_t^{\mathbf{Q}}\left[\frac{\frac{\hat{X}_T}{B_T}}{\frac{\hat{X}_t}{B_t}}\int_t^T \mathcal{D}_t \theta_s' dW_s^{\mathbf{Q}}\right] = \frac{\mathbf{E}_t\left[\xi_{t,T}^{1-1/R}\int_t^T (dW_s^{\mathbf{Q}})'\mathcal{D}_t\theta_s\right]}{\mathbf{E}_t\left[\xi_{t,T}^{1-1/R}\right]} \tag{2.16}$$

and $\xi_{t,T} = \exp(-\int_t^T (r_v + \frac{1}{2}\theta_v^2)dv - \int_t^T \theta_v dW_v)$. The Malliavin derivatives in a and b are given in theorem 34.

In this formula two expressions are provided for the coefficients a, b in the hedging components. The first is simply the specialization of the previous result to the case under consideration. The second formula uses the relation between optimal wealth and state prices in order to express a, b in terms of the relative state price density $\xi_{t,T}$ between periods t and T. This formula clearly demonstrates that the functions a, b depend only on the state variables Y.

The formulas described in theorems 34 and 26 provide useful information about the qualitative behavior of the investor. In order to assess the magnitude of the various components, and hence their relevance for asset allocation purposes, it is nevertheless necessary to get quantitative estimates. Practical implementations require the computation of the conditional expectations appearing in the portfolio formulas. Clearly Monte Carlo simulation appears to be an appealing way to proceed. In the next section we pursue this avenue and suggest a further transformation which facilitates the computation of Malliavin derivatives and may also help in the estimation of the hedging demands.

2.4 An alternative formula for Malliavin derivatives of diffusions

The key to our simplification is a change of variables which transforms a stochastic differential equation into an ordinary differential equation. In effect this (variance-stabilizing) transformation removes stochastic integrals from expressions such as $a(t, Y_t)$ and $b(t, Y_t)$. Changes of variables of this type are used by Doss (1977) to prove that an SDE can be solved pathwise, since it can be reduced to an ordinary differential equation.[9] Appendix A shows how Doss' arguments can be

[9] This result also plays an important role in the approximation of solutions of SDEs (e.g. Talay and Pardoux (1985)). In this context it can be used to conclude that convergence of the

used to derive alternative expressions for Malliavin derivatives of solutions of one-dimensional SDEs. In this section we state the result and discuss its implications.

2.4.1 The main result.

Consider a process Y which satisfies the one-dimensional SDE

$$dY_t = \mu(t, Y_t)dt + \sigma(t, Y_t)dW_t; Y_0 = y.$$

The Malliavin derivative of Y has the following alternative representation.

Proposition 27 *If the following conditions hold*[10]

(i) differentiability of drift: $\mu \in C^1([0,T] \times \mathbb{R})$

(ii) differentiability of volatility: $\sigma \in C^2([0,T] \times \mathbb{R})$

(iii) growth condition: $\mu(t,0)$ *and* $\sigma(t,0)$ *are bounded for all* $t \in [0,T]$,

then we have for $t \leq s$ *that*

$$\mathcal{D}_t Y_s = \sigma(s, Y_s) \exp\left[\int_t^s [\partial_2\mu - \frac{\mu\partial_2\sigma}{\sigma} - \frac{1}{2}(\partial_{22}\sigma)\sigma - \frac{\partial_1\sigma}{\sigma}](v, Y_v)dv\right] \tag{2.17}$$

Note that (2.17) expresses the Malliavin derivatives entirely in terms of Riemann-Stieltjes integrals of first and second derivatives of the coefficients of Y. Thus the stochastic integrals which appeared in the earlier formulas ((2.10) and (2.13)) have been entirely eliminated. Formula (2.17) is therefore easily computed using standard methods to approximate the Riemann integrals involved. With the variance stabilizing transformation the numerical calculation of the Malliavin derivatives is therefore of the same complexity as the numerical solution of an ODE. A second difference with the earlier expressions is that the leading term is the future

underlying Wiener process implies the convergence of the solution of an SDE.

[10]The space $C^i([0,T] \times \mathbb{R})$ is the space of i times continuously differentiable functions on the domain $[0,T] \times \mathbb{R}$.

volatility of the process at date s instead of the current volatility at t. This implies that this leading term cannot be factored out of conditional expectations at date t as was the case in (2.5) or (2.9). Randomness of the leading term however does not increase the computational difficulty involved in evaluating the Malliavin derivative.

With this numerically appealing expression for the Malliavin derivative we obtain a formula for the IR-hedge which does not involve stochastic integrals any longer. To achieve the same result for the MPR-hedge we introduce a second transformation which enables us to write the SPD and, as a consequence, also the MPR-hedge without any stochastic integral. We illustrate the idea in the univariate case.

Proposition 28 *Let $d = 1$. If the following conditions hold*

(i) differentiability of MPR: $\theta \in \mathcal{C}^2([0,T] \times \mathbb{R})$

(ii) differentiability of volatility: $\sigma \in \mathcal{C}^2([0,T] \times \mathbb{R})$

then the SPD can be written as

$$\xi_t = \exp\left[-\int_0^t [r + \frac{1}{2}\theta^2 - \frac{\theta}{\sigma}\mu - \partial_1\psi - \frac{1}{2}(\partial_2\theta\sigma - \theta\partial_2\sigma)](s, Y_s)ds - \psi(t, Y_t) + \psi(0, Y_0)\right] \tag{2.18}$$

where the function $\psi \in \mathcal{C}^1([0,T] \times \mathbb{R})$ solves $\partial_2\psi\sigma = \theta$. Consequently, we obtain

$$\int_t^T \mathcal{D}_t\theta_s[dW_s + \theta_s ds] = \frac{\theta}{\sigma}(T, Y_T)\mathcal{D}_tY_T - \theta(t, Y_t) - \int_t^T (g_1(s, Y_s) + g_2(s, Y_s))\, \mathcal{D}_tY_s ds \tag{2.19}$$

where

$$\begin{aligned} g_1(s, Y_s) &\equiv \left[\frac{\partial_1\theta}{\sigma} - \frac{\theta}{\sigma}\frac{\partial_1\sigma}{\sigma}\right](s, Y_s) \\ g_2(s, Y_s) &\equiv \left[\frac{1}{2}(\partial_{22}\theta\sigma - \theta\partial_{22}\sigma) + \frac{\partial_2\theta}{\sigma}\mu - \mu\frac{\theta}{\sigma}\frac{\partial_2\sigma}{\sigma} + \frac{\theta}{\sigma}\partial_2\mu - \theta\partial_2\theta\right](s, Y_s) \end{aligned}$$

2.4.2 A bivariate state variable example.

To illustrate the formulas above consider the model with CRRA of theorem 26 and suppose that the state variables are given by the pair (r, θ) which satisfies[11]

$$dr_t = \kappa_r(\overline{r} - r_t)dt + \sigma_r r_t^{\gamma_r} dW_t, \qquad r_0 \text{ given} \tag{2.20}$$

$$d\theta_t = \kappa_\theta(\overline{\theta} - \theta_t)dt + \sigma_\theta \theta_t^{\gamma_\theta} dW_t, \qquad \theta_0 \text{ given} \tag{2.21}$$

where $(\kappa_r, \overline{r}, \sigma_r, \gamma_r, \kappa_\theta, \overline{\theta}, \gamma_\theta)$ are nonnegative constants, $(\sigma_r, \sigma_\theta)$ are constants (possibly negative) and $(\gamma_r, \gamma_\theta) \in [0, 1]$. The Brownian motion W is unidimensional. This model nests standard formulations as special cases. The class of interest rate processes (2.20) is used in another context in Chan, Karolyi, Longstaff and Sanders (1992). The class of models (2.21) for the MPR has not been explored in the literature yet. We also assume that the stock volatility is stochastic and equal to $\sigma(r_t, \theta_t)$. This financial market is then described by two state variables (r, θ).

The transition from the general model with state variables Y to the model (2.20)-(2.21) with state variables (r, θ) is immediate since the Malliavin derivative $\mathbf{D}_t\theta_v$ can now be computed directly from the process (2.21). In order to state the result define the process

$$h_{t,v}(\gamma, \kappa, \sigma, \overline{x}; x) = -(1-\gamma) \int_t^v \left(\kappa(1 + \overline{x}\frac{\gamma}{1-\gamma}\frac{1}{x_u}) - \frac{1}{2}\sigma^2\gamma(\frac{1}{x_u})^{2(1-\gamma)} \right) du$$

for a quadruple of constants $(\gamma, \kappa, \sigma, \overline{x})$ and some process x. Taking account of the specific structure (2.20)-(2.21) then leads to

Corollary 29 *In the financial market (2.20)-(2.21) the optimal portfolio for CRRA utility is given by (2.14)-(2.16) where*

$$\mathbf{D}_t r_v = r_v^{\gamma_r}\sigma_r \exp\left[h_{t,v}(\gamma_r, \kappa_r, \sigma_r, \overline{r}; r)\right]$$

[11] This is equivalent to a model with two state variables $Y = (Y_1, Y_2)$ in which the equations $(r_t = r(t, Y_1), \theta_t = \theta\,(t, Y_2))$ can be inverted and the state variables can be expressed as $Y_t = (f_1(r_t), f_2(\theta_t))$.

$$\mathbf{D}_t\theta_v = \theta_v^{\gamma_\theta}\sigma_\theta \exp\left[h_{t,v}(\gamma_\theta,\kappa_\theta,\sigma_\theta,\bar{\theta};\theta)\right].$$

and

$$\xi_{t,T} = \exp\left[-\int_t^T r_s ds - \frac{1}{2}\int_t^T [\theta_s^2 + \sigma_\theta(1-\gamma_\theta)\theta_s^{\gamma_\theta} + 2\frac{\kappa_\theta}{\sigma_\theta}\theta_s^{1-\gamma_\theta}(\bar{\theta}-\theta_s)]ds - \phi(\theta_T) + \phi(\theta_t)\right]$$

with $\phi(x) = \frac{1}{\sigma_\theta(2-\gamma_\theta)}x^{2-\gamma_\theta}$. *The stochastic integral in the MPR-hedge (2.16) can also be written*

$$\int_t^T \mathcal{D}_t\theta_s[dW_s + \theta_s ds] = \theta_T \exp\left[h_{t,v}(\gamma_\theta,\kappa_\theta,\sigma_\theta,\bar{\theta};\theta)\right] - \theta_t - \int_t^T g_2(s,Y_s)\mathcal{D}_t\theta_s ds$$

with $g_2(s,Y_s) = \frac{1}{2}\sigma_\theta\gamma_\theta(1-\gamma_\theta)\theta_s^{\gamma_\theta-1} + \frac{\kappa_\theta}{\sigma_\theta}\left((1-\gamma_\theta)\bar{\theta}\theta_s^{-\gamma_\theta} - (2-\gamma_\theta)\theta_s^{1-\gamma_\theta}\right)$.

When $\gamma_r, \gamma_\theta = 0$ (Ornstein-Uhlenbeck IR and MPR processes) the formulas above simplify even further.

Corollary 30 *Suppose that* $u \in CRRA$. *When the interest rate and the market price of risk follow Ornstein-Uhlenbeck processes* ($\gamma_r, \gamma_\theta = 0$) *the optimal portfolio is given by (2.14)-(2.16) where*

$$a(t,r_t) = \frac{\sigma_r}{\kappa_r}\left(1 - \exp[-\kappa_r(T-t)]\right) \tag{2.22}$$

$$b(t,\theta_t) = \sigma_\theta \frac{\mathbf{E}_t\left[\xi_{t,T}^{1-1/R}\left(\int_t^T e^{-\kappa_\theta(v-t)}W_v^{\mathbf{Q}}\right)\right]}{\mathbf{E}_t\left[\xi_{t,T}^{1-1/R}\right]}. \tag{2.23}$$

The analytical expression for the IR-hedge in (2.22) clarifies the influence of the parameters of the interest rate process and the time horizon. Given the expressions provided in corollary 29 the MPR-hedge in (2.23) also has an analytical expression, albeit more complicated than the interest rate expression.

2.5 Numerical implementation

It follows from the results in the prior sections that the problem of finding the optimal portfolio for power utility function reduces to the identification of the functions a and b. When closed form expressions for a, b are not available, one must resort to a numerical scheme to estimate their values. As explained before Monte-Carlo simulation is naturally suggested by the structure of the problem and this is the approach that we adopt. In our context the simulation procedure involves two sources of error. First, since the joint law of the SPD and the Malliavin derivatives involved in the IR- and MPR- hedge terms are generally unknown we have to use a discretization scheme to approximate these random variables. It is well-known that such a discretization procedure produces a bias. Second, since we do not know how to calculate analytically the conditional expectation we rely on a law of large numbers to evaluate the expectation using independent replications of the random variables which enter in the hedging terms. This Monte-Carlo estimation of the conditional expectation also introduces an error.

In the discussion which follows we shall restrict attention to the model with CRRA utility. In this context, we estimate the functions a and b with M replications and N discretization points for the investment horizon by

$$a^{N,M}(T-t,y) = \frac{\sum_{i=1}^{M} (\xi_{T-t}^{N,i}(Y^N(y)))^{1-1/R} H_{T-t}^{a,N,i}(Y^N(y))}{\sum_{i=1}^{M} (\xi_{T-t}^{N,i}(Y^N(y)))^{1-1/R}}. \tag{2.24}$$

$$b^{N,M}(T-t,y) = \frac{\sum_{i=1}^{M} (\xi_{T-t}^{N,i}(Y^N(x))^{1-1/R} H_{T-t}^{b,N,i}(Y^N(y))}{\sum_{i=1}^{M} (\xi_{T-t}^{N,i}(Y^N(x))^{1-1/R}} \tag{2.25}$$

where $H_{T-t}^{a,N,i}(Y^N(y))$ and $H_{T-t}^{b,N,i}(Y^N(y))$ are estimators of $\int_t^T \mathcal{D}_t r_s ds$ and $\int_t^T \mathcal{D}_t \theta_s [dW_s + \theta_s ds]$ respectively. In these expressions we have emphasized that these quantities are functionals of the approximated state variables starting at $Y_0^N = y$.

Since the state variables, the SPD and the Malliavin derivatives of the state variables are all given as solutions of SDEs, the simplest approach for estimation is to use the Euler scheme. It has been shown by Kurtz and Protter (1991) that the order of convergence for this scheme is $1/\sqrt{N}$[12] due to the discretization error in the martingale parts of the SDEs. In Detemple, Garcia and Rindisbacher (DGR) (2000) we show that our variance-stabilizing transformation eliminates discretization errors in the martingale part of the SDE of the transformed state variables and therefore attains a rate of convergence of order $1/N$, which is the same convergence rate as for the Euler scheme of an ODE (see Appendix D).[13] In order to illustrate this difference in performance between the two schemes we estimate the respective absolute computational errors in the Malliavin derivative of the IR for different discretizations N of the time interval $[0, T]$. We estimate errors by the strong criterion

$$\widehat{\epsilon}^{N,M} = \hat{E}^M \left|\mathcal{D}_0^N r_T - \mathcal{D}_0 r_T\right| = \frac{1}{M}\sum_{i=1}^{M}\left|\mathcal{D}_0^{N,i} r_T - \mathcal{D}_0^i r_T\right|$$

where $\mathcal{D}_0 r_T$ denotes the true value of the derivative and $\mathcal{D}_0^N r_T$ its approximation based on N discretization points using M independent replications. We also compute the respective errors with and without transformation for the state variable r_T. Since the computation of this statistic requires the true distribution of the Malliavin derivative we assume that the IR follows the MRSR process ((2.20) with $\gamma_r = \frac{1}{2}$) with parameters $T = 1$, $\kappa_r = .004, \bar{r} = .06$, $\sigma_r = .0309839$, $r_0 = .06$.[14]

[12]That is $\sqrt{N}(Y^N - Y) \Rightarrow U^Y$ where convergence is in the weak sense and the error process in non-trivial $U^Y (\neq 0)$.

[13]That is $N(G(Z^N) - Y) \Rightarrow V^Y$, where $G(Z^N)$ is an estimator of the state variables Y and Z^N is obtained using the Euler scheme for the transformed state variables.

[14]Since $\sigma_r = 2\sqrt{\kappa_r \bar{r}}$ the interest rate r is the square of an Ornstein-Uhlenbeck process $Y_t = \sqrt{r_t}$. The true value can then be calculated by using the exact simulation of the transformed state variables

$$Y_{t+\Delta} = Y_t e^{\alpha\Delta} + \beta(\sigma_r e^{\alpha\Delta}\sqrt{\Delta}(W_{t+\Delta} - W_t) + \sqrt{|s_{22}|}Z)$$

where Z is a Gaussian variate independent of W, $\alpha = -\frac{\kappa_r}{2}$, $\beta = \sigma_r/2, \Delta = \frac{T}{N}$ and $s_{22} = e^{2\alpha\Delta}(\frac{1}{2\alpha} - \Delta) + 2(\Delta - \frac{1}{\alpha}) + \frac{3}{2\alpha}$. This choice of coefficients ensures that Y has the correct

To compute the expectation above we take 20 batches of 1,000 simulations each. For each batch an absolute error is estimated. Estimated absolute errors are then averaged over the batches. Table 1 below reports the results. Columns 2 and 4 show that the speed of convergence of the Euler scheme is roughly of order $1/\sqrt{N}$. Columns 3 and 5 illustrate the increase in the speed of convergence to $1/N$ when the scheme with transformation is used.

[Insert Table 1 here].

However, to compute hedging terms, we evaluate expectations of functionals of the state variables. In DGR (2000), we show that the increased speed of convergence obtained with the transformation for the numerical solution of SDEs of state variables fails to increase the speed of convergence of expectations of functionals of the state variables. This extends a result of Talay and Tubaro (1991). They have shown that, for the Euler scheme, $\mathbf{E}[f(Y_T^N) - f(Y_T)]$ is of order $\frac{1}{N}$ for functions f and diffusion coefficients μ and σ satisfying certain boundedness assumptions. Even though our problem is more complicated since we are not evaluating a function of a terminal point of a numerical solution to a SDE but a functional which depends on the whole trajectory of the solution, the same result holds. Nevertheless, as we will discuss next, the transformation may still be useful as it may reduce the asymptotic second order bias.

Denoting estimators without our transformation (by direct application of the Euler scheme) by $\tilde{\cdot}$ and estimators with the transformation by $\hat{\cdot}$, we obtain under certain integrability conditions (see DGR (2000) for details) for the $a(\cdot)$ function

$$\sqrt{M}(\tilde{a}^{N,M}(T-t,y) - a(T-t,y)) \Rightarrow \epsilon \tilde{K}_{T-t}^{a(y)} + M_{T-t}^{a(y)} \qquad (2.26)$$

variance and covariance with the increment of the Brownian motion $W_{t+\Delta} - W_t$.

$$\sqrt{M}(\hat{a}^{N,M}(T-t,y) - a(T-t,y)) \Rightarrow \epsilon \hat{K}_{T-t}^{a(y)} + M_{T-t}^{a(y)} \tag{2.27}$$

where $\epsilon = \frac{\sqrt{M}}{N}$ is fixed for all M, N. Corresponding limit laws are also obtained for the $b(\cdot)$ function. The vector processes $\tilde{K}^{a(y)}$ and $\hat{K}_{T-t}^{a(y)}$ (resp. $\tilde{K}^{b(y)}$ and $\hat{K}_{T-t}^{b(y)}$) are deterministic whereas $M^{a(y)}$ (resp. $M^{b(y)}$) is a Gaussian martingale. As indicated both types of processes depend on the initial position of the state variables, y.

In these expressions, the deterministic processes $K^{\cdot}$ correspond to the discretization error resulting from the approximation scheme and therefore depend on the approximation method used. Ideally, they should be zero. Using our transformation this is indeed the case if the underlying state variables are given by an invertible, twice continuously differentiable function of lognormal processes. It happens in this case that the approximation using the transformation is also exact for the part of the SDE involving Riemann integrals. But in general $\hat{K}^{\cdot}$ will be different from zero. Therefore, although the estimators are consistent, a smaller $\hat{K}^{\cdot}$ reduces the second order bias. If in the construction of confidence intervals we do not correct for this second order bias the size distortion[15] will be smaller with the transformation whenever $\hat{K}^{\cdot} < \tilde{K}^{\cdot}$. Consequently, a reduced second order bias will also improve the validity of statistical tests based on the law of $M^{\cdot}$ only. Furthermore, a small second order bias is potentially important for a good performance of the estimators given a finite number of replications and discretization points.

The processes $M^{\cdot}$ are for both approximation methods the same. They result from the Monte Carlo estimation of the conditional expectation and would not vanish even if we could sample from the true joint law of $H^{\cdot}$ and the SPD ξ. The expressions for both processes K and M are obtained in explicit form and

[15]Size distortion refers to the fact that the actual coverage probability is different from the prescribed level.

described in detail in DGR (2000) and can therefore be used to implement error corrections and variance reductions.

All the results discussed above are conditional on the knowledge of the state variables at a given moment in time. If we are only interested at point estimators of the optimal composition of our portfolio given a certain state the estimators of a and b are all we have to calculate. But for other purposes, such as risk management, we may well be interested in testing a given portfolio strategy against a specific benchmark. Since this type of exercise requires the probabilistic structure of the optimal portfolio strategy, we need the distribution of conditional estimators of the mean-variance component, the IR-hedge and the MPR-hedge. Since we cannot sample from the true law of the state variables it follows that we have to rely on an approximation of their dynamic evolution described by the SDE. As we show in DGR (2000) the conditional estimators converge weakly with order $\frac{1}{N}$ with transformation and order $\frac{1}{\sqrt{N}}$ without. The limit laws of these conditional estimators are non-Gaussian [16] but known and therefore can be used to construct asymptotically valid confidence intervals or statistical tests.

2.6 Calibration of the model

In order to examine the economic properties of optimal portfolios we need to specify and calibrate our model of the financial market. We will focus on the class of bivariate processes for (r, θ) described in the section above. Specifically we estimate the following IR-MPR model

$$dr_t = \kappa_r(\overline{r} - r_t)dt - \sigma_r r_t^{1/2} dW_t, \qquad r_0 \text{ given} \tag{2.28}$$

$$d\theta_t = \kappa_\theta(\overline{\theta} - \theta_t)dt + \sigma_\theta dW_t, \qquad \theta_0 \text{ given} \tag{2.29}$$

[16] The reader is referred to DGR (2000) for the exact expressions.

where $(\kappa_r, \overline{r}, \sigma_r, \gamma_r, \kappa_\theta, \overline{\theta}, \sigma_\theta, \gamma_\theta)$ are constants.

We assume that the approximate discrete-time process is the true time-series model.[17] The econometric procedure described in this section is based on the maximization of the loglikelihood of the following discrete-time model:

$$r_{t_{n+1}}^{(h)} = r_{t_n}^{(h)} + \kappa_r(\overline{r}_h - r_{t_n}^{(h)}) + \sigma_{r,h}\sqrt{r_{t_n}^{(h)}}\varepsilon, \qquad r_0 \text{ given} \tag{2.30}$$

$$\theta_{t_{n+1}} = \theta_{t_n} + \kappa_\theta(\overline{\theta} - \theta_{t_n}) + \sigma_\theta v, \qquad \theta_0 \text{ given.} \tag{2.31}$$

where $\overline{r}_h = \overline{r}h$ and $\sigma_{r,h} = \sigma_r\sqrt{h}$ and $\{t_n : n = 0, ..., N\}$ is a partition of $[0, T]$. In our estimations, we consider a monthly frequency with $h = \frac{1}{12}$.

Since the MPR, $\theta_t = \sigma_t^{-1}(\mu_t - r_t)$, is unobservable it must be filtered from the data. We take two approaches.[18] First we assume that the stock volatility σ is constant. In other words, we estimate the MPR from the conditional mean μ_t of the stock return series (taken as the SP500 index), assuming a simple AR(1) process for the conditional mean. The estimation period is January 1965-June 1996.

In the continuous-time model the same Brownian motion applies to r and θ, but with a perfect negative correlation. We therefore produce two sets of estimates, one with the correlation coefficient between ε_{t+1} and v_{t+1} left unconstrained, an-

[17]Estimating the parameters of a continuous-time diffusion model based on a discrete-time approximation of the likelihood function leads to a discretization bias (Lo (1988)). However, for the monthly estimation of interest rate processes, Broze, Scaillet and Zakoian (1995) use an indirect estimation to correct for the bias and find that the bias is small for the mean-reversion κ_r, the mean $\overline{r}$ and the variance σ_r. We therefore follow the simpler approach to calibrate the parameters. We also investigate the sensitivity of the results to changes in the parameters.

[18]This filtering approach is in the spirit of Nelson and Foster (1994), although we do not claim any optimality property for the GARCH(1,1) process we use.

other one with a negative correlation of -0.9.[19] The results are presented in Tables 2 and 3 respectively. The estimates obtained for the parameters of the interest rate CIR process are comparable to the values obtained by Broze, Scaillet and Zakoïan (1995) and Chan et al. (1992). The process slowly reverts to an annualized mean of about 6% with a yearly volatility of about 1.76% for the unconstrained model and around 3.6% for the constrained estimate. The estimation results for the MPR Orstein-Uhlenbeck process show that the market price of risk reverts rather quickly to its mean. The mean is about 8%, which is low compared with the standard estimates of the market price of risk. The MPR volatility is about ten times the volatility of the interest rate process in both the unconstrained and constrained estimations; almost perfect negative correlation between the interest rate and the MPR forces upwards the volatilities of the two processes by a factor of two. Given the low value of the MPR, we also investigate a specification where the interest rate enters in the drift of the market price of risk, since excess returns are known to be predictable by the interest rate. Equation (2.32) replaces (2.31)

$$\theta_{t_{n+1}} = \theta_{t_n} + \kappa_\theta(\overline{\theta} - \theta_{t_n}) + \delta r_{t_n}^{(h)} + \sigma_\theta v, \qquad \theta_0 \text{ given.} \tag{2.32}$$

The estimation results, reported in Table 4, are quite similar to the previous specification, except for the mean level of the MPR, which is more in line with the usual estimate of 0.3. The expected negative coefficient of the interest rate δ comes out quite significantly different from zero. As we will see, this specification will only change the absolute magnitude of the stock position and the hedging terms, but not the relative importance of the later with respect to the former.

To assess the robustness of the results obtained with a constant σ,we use a GARCH (1,1) model for the stock returns to construct the series for the market

[19] Since at a correlation of -1, the variance-covariance matrix would be singular, we chose the closest approximation that did not create numerical problems.

price of risk θ_t. We keep as before an AR(1) specification for the conditional mean of the stock returns. The results are reported in Tables 5 and 6, where as before we estimate two versions of the model, with correlation coefficient $\rho_{r\theta}$ left unconstrained (Table 5) and constrained to a value of -0.9 (Table 6). The most notable differences are a moderate increase (decrease) in the interest rate (MPR) speed of mean-reversion by about 10%, an increase in the long run mean of the MPR by about 5% and a decrease in the MPR volatility by about 7%. The estimates obtained for the other parameters are roughly the same as before. Overall these differences should not exert much influence on the magnitude of the hedging terms and will not be considered in our numerical computation of optimal portfolios.

2.7 Economic properties of optimal portfolios

We now implement our numerical procedure for the model with (i) constant relative risk aversion, (ii) a single risky stock with constant volatility, (iii) an MRSR (mean reverting - square root) process for the interest rate and (iv) a MRG (mean-reverting Gaussian) process or a MRGID (mean-reverting with interest rate dependence in the drift) process for the MPR. The uncertainty is thus captured by a bivariate system of state variables (r, θ). For this specification of preferences and uncertainty we recall that the stock demand is

$$\hat{\pi}_t = \hat{X}_t \frac{1}{R} \sigma^{-1} \theta_t + \hat{X}_t \left(\frac{1}{R} - 1 \right) \sigma^{-1} a(t, r_t) + \hat{X}_t \left(\frac{1}{R} - 1 \right) \sigma^{-1} b(t, \theta_t) \tag{2.33}$$

$$a(t, r_t) := \frac{\mathbf{E}_t \left[\xi_{t,T}^{1-1/R} \int_t^T \mathcal{D}_t r_s ds \right]}{\mathbf{E}_t \left[\xi_{t,T}^{1-1/R} \right]}, \tag{2.34}$$

$$b(t, \theta_t) := \frac{\mathbf{E}_t \left[\xi_{t,T}^{1-1/R} \int_t^T \mathcal{D}_t \theta_s dW_s^{\mathbf{Q}} \right]}{\mathbf{E}_t \left[\xi_{t,T}^{1-1/R} \right]} \tag{2.35}$$

where $\mathcal{D}_t r_s, \mathcal{D}_t \theta_s$ and $\xi_{t,T}$ are provided in corollary 29. For the MRGID model see Appendix C.

Parameter values are set at their estimated values reported in Tables 3 and 4 and at values equal or close to the means for r_0 and θ_0; the volatility of the stock is set at its historical average 0.2. Specifically, in the first model (Table 3), we take $\kappa_r = .0824$, $\overline{r} = .0050 \times 12$, $\gamma_r = .5$, $\sigma_r = .01050 \times \sqrt{12}$ (recall that there is a minus sign in front of σ_r in (2.28)), $r_0 = .0050 \times 12$, $\sigma = .20$, $\kappa_\theta = .6950$, $\overline{\theta} = .0871$, $\gamma_\theta = 0$, $\sigma_\theta = .21$, $\theta_0 = .10$. In the second model, we take the following values: $\kappa_r = .0005$, $\overline{r} = .0050 \times 12$, $\gamma_r = .5$, $\sigma_r = .01050 \times \sqrt{12}$, $r_0 = .0050 \times 12$, $\sigma = .20$, $\kappa_\theta = .7771$, $\overline{\theta} = .2675$, $\gamma_\theta = 0$, $\sigma_\theta = .205, \delta = -26.29/12$, $\theta_0 = .30$. Simulations are carried out using daily increments and $5,000$ paths with variance reduction by antithetic variables method ($M = 5,000$, $h = 1/365$). Since the results are very similar, except for the difference in the absolute magnitude of the hedging terms as mentioned earlier, we only report in Table 7 summary results for the first model and provide a full-fledged analysis with graphs for the second model.[20]

2.7.1 Optimal portfolios and hedging components.

Figures 1-3 illustrate the behavior of the optimal portfolio and the hedging components relative to the risk aversion coefficient and the investment horizon. Risk aversion varies from .5 to 5; the investment horizon from 1 year to 5 years. As expected the fraction of wealth invested in the stock decreases as risk aversion increases and increases as the horizon increases. The hedges, however, display strikingly different behavior. The MPR-hedge displays mildly humped decreasing-

[20] Of course the expressions for the Malliavin derivatives with respect to the interest rate and the MPR as well as the state price density change compared to our previous expressions in section 5. The new expressions are given in Appendix C.

increasing behavior relative to risk aversion and appears to decrease relative to horizon. The IR-hedge increases relative to both variables. As noted before the signs of the hedges change depending on whether risk aversion exceeds or falls short of 1. This illustrates the standard knife-edge behavior of (myopic) logarithmic utility. For investors that are more risk averse than the Bernoulli investor the negative values of the MPR-hedge stem from the positive correlation between the stock return and the MPR. Such an investor tries to hedge the additional risk away by reducing his/her stock demand. Similarly the IR-hedge tends to boost stock demand since it covaries negatively with the stock return. Note also that the combination of the two hedges is negative for short investment horizons (less than 4 years in the numerical example) and positive for longer holding periods. Thus, hedging behavior reduces (increases) the stock investment for short run (long run) horizons relative to a pure mean-variance investor. In fact, the increase in stock holdings increases with longer investment horizons.

[Insert figures 1-3 here]

Figures 4-6 display the behavior relative to the levels of the IR and the MPR r_0, θ_0 for risk aversion $R = 2$ and investment horizon $T - t = 1$. Again the fraction invested in the stock varies considerably over the range of initial values investigated, from over 90% of wealth to nearly 25%. The hedge components' ranges are much narrower: while the IR-hedge varies between about 1.8% and 2.6%, the MPR-hedge lies between -1.8% and about -6%.

Second note that the fraction invested in the stock is an increasing function of the MPR and is almost insensitive to the interest rate. As θ_0 increases the IR-hedge stays flat (figure 5) while the MPR-hedge becomes more negative (figure 6). These effects, however, are of second order relative to the increase in the mean-variance component of the stock demand. When r_0 increases the IR-hedge increases moderately and becomes more positive (see figure 6): it tends to increase stock demand. The MPR-hedge also increases but even more moderately. Combining these two effects produces a mildly increasing total stock demand. For typical values of the MPR (between .20 and .40) the sum of the hedging terms is negative and tends to reduce the overall demand for the stock.

[Insert figures 4-6 here].

2.7.2 Market timing strategies.

In order to assess the importance and stability over time of our hedging demand estimates we perform two market timing experiments. The first consists in drawing trajectories of the underlying state variable processes r, θ and computing the portfolio and hedging demands along these trajectories. The second experiment simulates the optimal portfolio for very long horizons and using actual market data.

Results for the first experiment are reported in figures 7-10. A typical trajectory of the pair (r, θ) is drawn in figures 7 and 8. The interest rate is seen to vary between 3.9% and 5.4%; the MPR takes values between $-.08$ and .30. Figure 9 illustrates the stock demand behavior for an investor with risk aversion of 4 and a fixed horizon of 5 years. For the trajectory drawn the proportion invested in the stock evolves between -3% and 40%. Close inspection of the graph, however, shows that changes superior to 30% in the portfolio share are usually spread over periods of 6 month or more. There are also long stretches of time, of duration larger than a year, over which the stock share varies within at 10% interval.

Figure 10 which shows the respective contributions of the IR-hedge, the MPR-hedge and the sum of the two hedges sheds further light on this issue. First note that the IR-hedge is remarkably stable over time. It experiences very small fluctuations and decreases slowly toward zero due to the maturity effect of the fixed horizon. It also remains positive throughout the period. The MPR-hedge is negative and exhibits stronger volatility, which is not surprising since it is sensitive to the MPR level which is more volatile. Within intervals of a year though the fluctuations rarely exceed 5%. Again a trend toward zero is observed due to the fixed investment horizon. Both hedges work in opposite direction and partly offset each other. The net hedging correction is of the order of $5\% - 10\%$ at the beginning of the investment horizon, thus boosting the stock demand. It then slowly converges toward zero taking negative values along the way, thus reducing stock demand, in the last couple of years of the period. The net hedging correction inherits the stability of its two components: its fluctuations rarely exceeds 5% over periods of a year or longer. Over the whole 5 year period the hedging correction varies between -3% and 10%.

Although not reported in the paper similar properties are recorded when the analysis is performed for rolling horizons of 2 years and 5 years (though hedging terms do not converge to zero in that case) and for risk aversions in the range $2 - 4$.

We conclude from this (representative) experiment that hedging components are remarkably stable over time in the sense that they exhibit low volatility. The variation in the total stock demand which is observed in figure 9 stems primarily from the variation of its mean variance component.

[Insert figures 7-10 here]

Our second experiment examines the actual behavior, based on market data, of the portfolio over time for an investor with long horizon of about 30 years at the beginning of the period. Hedging demands and portfolio positions are computed using our model along the realized trajectory of the IR and the MPR in the last 31.5 years (our estimation sample). Based on these data, we compute each month of the sample the optimal share of the stock in the portfolio with and without hedging for an investor with a relative risk aversion of 4 (computations are performed using 25,000 replications and variance reduction, i. e. 50,000 replications). As figure 11 shows, intertemporal hedging will increase the optimal share to a reasonable level of about 60% at the beginning of the investment horizon to roughly 10% at the end, with an average holding of 44%. This is in sharp contrast with the myopic mean-variance optimal share which varies substantially around an average level of about 10%. Note also that the hedging investor will short the stock by 15% only once during the investment period (during the 1987 crash) and only because the triggering event happened shortly (10 years) before the end of the investment horizon. The observed increase in stock holdings comes mainly from the positive IR-hedge. From this realistic situation we then conclude that intertemporal hedging has a fundamental impact when the investment horizon is long. As in the previous experiment it tends to stabilize the overall stock demand.

[Insert figure 11 here]

2.8 Stochastic dividends (trivariate model)

Suppose now that the dividend-price ratio (DPR), denoted by p, is a relevant stochastic factor which influences the evolution of the market price of risk. The following trivariate process for (r, θ, p) generalizes the MRGID model by incorporating such an effect

$$dr_t = \kappa_r(\overline{r} - r_t)dt - \sigma_r r_t^{1/2} dW_t, \qquad r_0 \text{ given} \tag{2.36}$$

$$d\theta_t = [\kappa_\theta(\overline{\theta} - \theta_t) + \delta_r r_t + \delta_p p_t]dt + \sigma_\theta dW_t, \qquad \theta_0 \text{ given} \tag{2.37}$$

$$dp_t = \kappa_p(\overline{p} - p_t)dt - \sigma_p p_t^{1/2} dW_t, \qquad r_0 \text{ given.} \tag{2.38}$$

In this specification the DPR follows a mean-reverting square root process and has a linear effect on the drift of the MPR.

The model is estimated as previously: we maximize the loglikelihood of the discretized model using for the MPR the filtered series based on an AR(1) specification and a constant stock volatility. For the sake of brevity, we just report the estimated values of the parameters. These are $\kappa_r = 0.06977$, $\overline{r} = 0.005 \times 12$, $\kappa_\theta = 0.9088$, $\overline{\theta} = 0.1685$, $\delta_r = -23.90/12$, $\delta_p = 17.63/12$, $\kappa_p = 0.0344$, $\overline{p} = 0.003 \times 12$, $\sigma_r = 0.01227\sqrt{12}$, $\sigma_\theta = 0.16127$, $\sigma_p = 0.004578\sqrt{12}$. It should, however, be noted that these estimates, in particular those corresponding to the impact of the IR and the DPR on the drift of the MPR, are statistically different from zero. Other parameters are also seen to be close to the values obtained for the model with two state variables only.

Table 8 shows that optimal behavior changes when stochastic dividends are accounted for. The most notable feature is the reversal in the sign of the MPR-hedge. Inspection of the trivariate process reveals the root of this behavior. Recall that the estimated model displays positive impact of the dividend-price ratio on the drift of the MPR ($\delta_p = 17.63/12$) and negative correlation between stock returns and the dividend-price ratio ($-\sigma\sigma_p = -0.2 \times 0.004578\sqrt{12}$). Under these conditions hedging MPR-risk will involve two components. The first results from the positive association between stock returns and innovations in the MPR. This hedge against direct MPR-risk is negative, as in the earlier models. The second is the consequence of the indirect negative association between the drift of the MPR and innovations in the dividend-price ratio. This hedge, against indirect MPR-risk, is positive. Evidently, the two hedging motives work in opposite direction.

As illustrated in the table (see also figure 12) the second effect dominates in the context of our estimated model and results in a positive overall MPR-hedge when risk aversion exceeds unity.

2.9 Stochastic volatility with imperfect correlation (trivariate model)

Consider now the trivariate state variable model (r, θ, σ) described by

$$dr_t = \kappa_r(\overline{r} - r_t)dt - \sigma_r r_t^{1/2} dW_{1t} \tag{2.39}$$

$$d\theta_t = [\kappa_\theta(\overline{\theta} - \theta_t) + \delta r_t]dt + \sigma_\theta dW_{1t} \tag{2.40}$$

$$d\sigma_t = [\kappa_\sigma(\overline{\sigma} - \sigma_t) + \sigma_t\{\delta_{1\theta}\theta_t \mathbf{1}_{\{\theta \geq 0\}} + \delta_{2\theta}\theta_t \mathbf{1}_{\{\theta < 0\}}\}]dt + \sigma_t^{1/2}[\lambda_1 dW_{1t} + \lambda_2 dW_{2t}] \tag{2.41}$$

where $(r_0, \theta_0, \sigma_0)$ are given, the coefficients $(\kappa_r, \overline{r}, \sigma_r, \kappa_\theta, \overline{\theta}, \delta, \sigma_\theta, \kappa_\sigma, \overline{\sigma}, \delta_{1\theta}, \delta_{2\theta}, \lambda_1, \lambda_2)$ are all constant and W_1, W_2 are independent Brownian motion processes.

The model (2.39)-(2.41) contains several innovations relative to the prior MRSR-MRGID model. The most important feature is that volatility is now stochastic. Furthermore, the volatility process is imperfectly correlated with the interest rate and the MPR processes. As a result our basic model is one with (apparently) incomplete markets. The drift of the volatility process also permits an asymmetric dependence on the MPR process, conditioned on positive or negative realizations of the MPR. This structure seeks to capture the notion that volatility is high when the magnitude (absolute value) of the MPR is large. As in the MRGID model the MPR process also involves an interaction in the drift with the rate of interest.

Even though this trivariate model (2.39)-(2.41) is driven by two underlying Brownian motions, and hence appears to have incomplete markets since there are only two assets, the portfolio formulas of the previous sections are still valid. The intuition for this seemingly surprising result is that the state price density ξ depends only on (r, θ) which are independent of the risk W_2. Since the investor's marginal utility is proportional to the state price density at the optimum it follows that optimal terminal wealth is independent of W_2. The portfolio that finances optimal wealth, in turn, will be independent of this idiosyncratic volatility risk. It follows that the individual valuation of the risk W_2 is null at the optimum.

Assuming CRRA preferences gives the optimal stock demand

$$\hat{\pi}_t = \hat{X}_t \frac{1}{R} \sigma_t^{-1} \theta_t + \hat{X}_t \left(\frac{1}{R} - 1 \right) \sigma_t^{-1} a(t, r_t, \theta_t) + \hat{X}_t \left(\frac{1}{R} - 1 \right) \sigma_t^{-1} b(t, r_t, \theta_t) \quad (2.42)$$

$$a(t, r_t, \theta_t) \equiv \frac{\mathbf{E}_t \left[\xi_{t,T}^{1-1/R} \int_t^T \mathcal{D}_t r_s ds \right]}{\mathbf{E}_t \left[\xi_{t,T}^{1-1/R} \right]}, \quad (2.43)$$

$$b(t, r_t, \theta_t) \equiv \frac{\mathbf{E}_t \left[\xi_{t,T}^{1-1/R} \int_t^T \mathcal{D}_t \theta_s dW_{1s}^{\mathbf{Q}} \right]}{\mathbf{E}_t \left[\xi_{t,T}^{1-1/R} \right]} \quad (2.44)$$

where $\xi_{t,T}$ is defined in corollary 29 and where $\mathcal{D}_t r_s, \mathcal{D}_t \theta_s$ are given in explicit form in appendix C. The only notable impact of stochastic volatility is that it implies a continuous rescaling of the stock demand as it changes over time: the volatility-scaled portfolio demand $\sigma_t \hat{\pi}_t$ is immune to volatility risk.

The economic properties of the optimal portfolio follow directly from the scaling property. The fraction of each hedging demand relative to total stock demand is insensitive to volatility fluctuations. Since the magnitude of each component is simply rescaled as volatility changes the portfolio components exhibit more volatility. This behavior is illustrated in figure 13.[21]

[21] Again for this extension of the basic two-state variable model, we maximize the loglikelihood of the discretized model. The MPR series is now filtered with a GARCH(1,1) model with

2.10 A multiasset-trivariate model: hedging with two mutual funds

We now consider a financial market with three assets (2 risky and a riskless asset) and a triplet of state variables. Our objectives are to provide a decomposition of the optimal portfolio and to examine the effects of correlations on the hedging terms.

The state variables (r, θ_1, θ_2) evolve according to

$$dr_t = \kappa_r(\bar{r} - r_t)dt - \sigma_r r_t^{1/2} dW_{1t}, \qquad r_0 \text{ given} \tag{2.45}$$

$$d\theta_{1t} = \left(\kappa_1(\bar{\theta}_1 - \theta_{1t}) + \delta_{1r} r_t\right) dt + \sigma_1^\theta \frac{1}{\sqrt{2\alpha}} \left(\alpha dW_{1t} + \rho_\theta dW_{2t}\right), \quad \theta_0^1 \text{ given} \tag{2.46}$$

$$d\theta_{2t} = \left(\kappa_2(\bar{\theta}_2 - \theta_{2t}) + \delta_{2r} r_t\right) dt + \sigma_2^\theta \frac{1}{\sqrt{2\alpha}} \left(\rho_\theta dW_{1t} + \alpha dW_{2t}\right), \quad \theta_0^2 \text{ given} \tag{2.47}$$

where $\alpha = 1 + \sqrt{1 - \rho_\theta^2}$ and $(\kappa_r, \bar{r}, \sigma_r, \kappa_1, \bar{\theta}_1, \delta_{1r}, \sigma_1^\theta, \kappa_2, \bar{\theta}_2, \delta_{2r}, \sigma_2^\theta, \rho_\theta)$ are constants. In this formulation σ_i^θ is the standard deviation of $\theta_i, i = 1, 2$, and ρ_θ represents the correlation coefficient between θ_1 and θ_2. The correlation between the interest rate and the market price of W_1-risk is negative and equals $\rho_{r\theta_1} = -\sqrt{\frac{1}{2}(1 + \sqrt{1 - \rho_\theta^2})}$. The correlation with the market price of W_2-risk, which equals $\rho_{r\theta_2} = -\rho_\theta / \sqrt{2(1 + \sqrt{1 - \rho_\theta^2})}$, is negative (positive) when ρ_θ is positive (negative). When MPRs are positively correlated (ρ_θ positive) an increase in their correlation will increase $\rho_{r\theta_1}$ and decrease $\rho_{r\theta_2}$.

an AR(1) conditional mean as described in section 6. The estimates of the parameters in (r, θ) are found to be stable relative to those obtained in the earlier bivariate model. For the volatility process we find evidence of different effects for the positive and negative values of the MPR. This confirms the asymmetry reported in the literature. Estimated parameters are $\kappa_r = 0.004575$, $\bar{r} = 0.007 \times 12$, $\kappa_\theta = 0.7772$, $\bar{\theta} = 0.2689$, $\delta = -26.3514/12$, $\kappa_\sigma = 0.0445$, $\bar{\sigma} = 0.0594\sqrt{12}$, $\delta_{1\theta} = -0.2159/12$, $\delta_{2\theta} = 0.1254/12$, $\sigma_r = 0.01045\sqrt{12}$, $\sigma_\theta = 0.185$, $\lambda_1 = 0.00081$, and $\lambda_2 = \sqrt{0.01252^2 + 0.00174^2}$. The numerical simulation is based on these estimates.

The riskless asset pays interest at the rate r in (2.45). The price S_i of asset i, $i = 1, 2$, satisfies

$$dS_{it} + \delta_{it} S_{it} dt = S_{it} \left[\mu_{it} dt + \sigma_i \left(\rho_i dW_{1t} + \sqrt{1 - \rho_i^2} dW_{2t} \right) \right] \tag{2.48}$$

where the dividend rate δ_i and the drift μ_i are stochastic. The volatility matrix of asset returns is constant and assumed to be invertible (i.e. $\Delta \equiv \sigma_1 \sigma_2 (\rho_1 \sqrt{1 - \rho_2^2} - \rho_2 \sqrt{1 - \rho_1^2}) \neq 0$). Asset prices induce the (bivariate) MPR process (θ_1, θ_2) whose evolution is described in (2.46)-(2.47). The first risky asset can be interpreted as the market portfolio of risky stocks (SP500); the second is a portfolio of assets (mutual fund) whose correlation with the market portfolio is $\rho = \rho_1 \rho_2 + \sqrt{1 - \rho_1^2} \sqrt{1 - \rho_2^2}$. The correlation coefficients between the portfolio returns and the interest rate are respectively $-\rho_1$ and $-\rho_2$.

In this setting with two assets the optimal portfolio is given by[22]

[22]The structure of the optimal portfolio remains the same if the volatility coefficients σ_1, σ_2 are stochastic. As in the prior section structure is preserved even when volatility risk can only be partially hedged with traded assets.

$$\begin{bmatrix} \hat{\pi}_{1t}/\hat{X}_t \\ \hat{\pi}_{2t}/\hat{X}_t \end{bmatrix} = \frac{1}{R}\frac{1}{\Delta}\begin{bmatrix} \sigma_2\sqrt{1-\rho_2^2}\theta_{1t} - \sigma_2\rho_2\theta_{2t} \\ -\sigma_1\sqrt{1-\rho_1^2}\theta_{1t} + \sigma_1\rho_1\theta_{2t} \end{bmatrix}$$
$$+(\frac{1}{R}-1)\frac{1}{\Delta}\begin{bmatrix} \sigma_2\sqrt{1-\rho_2^2} \\ -\sigma_1\sqrt{1-\rho_1^2} \end{bmatrix}\overline{a}(t,r_t,\theta_t)$$
$$+(\frac{1}{R}-1)\frac{1}{\Delta}\begin{bmatrix} \sigma_2\sqrt{1-\rho_2^2}\alpha - \rho_\theta\sigma_2\rho_2 \\ -\sigma_1\sqrt{1-\rho_1^2}\alpha + \rho_\theta\sigma_1\rho_1 \end{bmatrix}\sigma_1^\theta\frac{1}{\sqrt{2\alpha}}\overline{b}_1^{dir}(t,r_t,\theta_t)$$
$$+(\frac{1}{R}-1)\frac{1}{\Delta}\begin{bmatrix} \sigma_2\sqrt{1-\rho_2^2} \\ -\sigma_1\sqrt{1-\rho_1^2} \end{bmatrix}\delta_{1r}\overline{b}_1^{ind}(t,r_t,\theta_t)$$
$$+(\frac{1}{R}-1)\frac{1}{\Delta}\begin{bmatrix} \rho_\theta\sigma_2\sqrt{1-\rho_2^2} - \alpha\sigma_2\rho_2 \\ -\rho_\theta\sigma_1\sqrt{1-\rho_1^2} + \alpha\sigma_1\rho_1 \end{bmatrix}\sigma_2^\theta\frac{1}{\sqrt{2\alpha}}\overline{b}_2^{dir}(t,r_t,\theta_t)$$
$$+(\frac{1}{R}-1)\frac{1}{\Delta}\begin{bmatrix} \sigma_2\sqrt{1-\rho_2^2} \\ -\sigma_1\sqrt{1-\rho_1^2} \end{bmatrix}\delta_{2r}\overline{b}_2^{ind}(t,r_t,\theta_t) \qquad (2.49)$$

where

$$\overline{a}(t,r_t,\theta_t) = \mathbf{E}_t\left[\frac{\xi_{t,T}^{1-1/R}}{\mathbf{E}_t\left[\xi_{t,T}^{1-1/R}\right]}\int_t^T \mathcal{D}_{1t}r_s ds\right]$$

$$\overline{b}_i^{dir}(t,r_t,\theta_t) = \mathbf{E}_t\left[\frac{\xi_{t,T}^{1-1/R}}{\mathbf{E}_t\left[\xi_{t,T}^{1-1/R}\right]}\int_t^T e^{-\kappa_i(s-t)}dW_{is}^{\mathbf{Q}}\right], \quad i=1,2$$

$$\overline{b}_i^{ind}(t,r_t,\theta_t) = \mathbf{E}_t\left[\frac{\xi_{t,T}^{1-1/R}}{\mathbf{E}_t\left[\xi_{t,T}^{1-1/R}\right]}\int_t^T\int_t^s e^{-\kappa_i(s-v)}\mathcal{D}_{1t}r_v dv dW_{is}^{\mathbf{Q}}\right], \quad i=1,2.$$

and $\mathcal{D}_{1t}r_s = -\sigma_r r_s^{\frac{1}{2}}\exp(-\frac{1}{2}\int_t^s(\kappa_r - \frac{\sigma_r^2}{4})\frac{1}{r_s}ds - \frac{1}{2}\kappa_r(s-t))$ is nonpositive at all times (taking $\sigma_r > 0$). The first line in (2.49) is the MV-demand, the second the

IR-hedge, the third and fourth the MPR(θ_1)-hedge and the last two the MPR(θ_2)-hedge. The function $\overline{a}(t, r_t, \theta_t)$ is the cross-moment between the cost of optimal consumption $\xi_{t,T}^{1-1/R}$ and the sensitivity of the cumulative interest rate to W_1-risk (i.e. $\int_t^T \mathcal{D}_{1t} r_s ds$). Since θ_1 and θ_2 depend on the interest rate (see (2.46)-(2.47)), innovations in W_i will have a direct effect on future values of the MPRs as well as an indirect effect through the interest rate. The covariances $\overline{b}_i^{dir}(t, r_t, \theta_t)$ and $\overline{b}_i^{ind}(t, r_t, \theta_t)$ capture, respectively, these two aspects.

Let us focus on the effects of correlation between the two funds. Assume that all the coefficients are positive except for the correlation coefficients $\rho_1, \rho_2, \rho_\theta$ which may take positive or negative values. As it turns out the sign of all the demand components result from the spanning properties of the two traded assets and the risk exposure of the present value of terminal (optimal) consumption (PVC). This follows since the optimal portfolio is selected so as to finance this present value.

In particular, note that the MV components result from the desire to synthesize the vector (θ_1, θ_2) which describes the risk exposure of the state price density and captures the impact of the SPD on the PVC. When (θ_1, θ_2) is a convex combination of the vectors generated by the two funds returns, namely $(\rho_1, \sqrt{1-\rho_1^2})$ and $(\rho_2, \sqrt{1-\rho_2^2})$ then both demands are positive. This is the case when $\rho_1/\sqrt{1-\rho_1^2} > \theta_1/\theta_2 > \rho_2/\sqrt{1-\rho_2^2} > 0$. Otherwise, one fund is held short and the other long, and the MV demands are of opposite signs.

The IR-hedges in the two portfolio components reflect similar considerations. Here it is the risk exposure of the PVC induced by the interest rate that is being synthesized, i.e. the vector

$$(\frac{1}{R} - 1) \begin{bmatrix} 1 \\ 0 \end{bmatrix} \overline{a}(t, r_t, \theta_t).$$

Since this risk exposure is outside the convex cone generated by asset returns demands will necessarily be of opposite sign. When risk aversion exceeds one interest rate risk has a positive impact on the PVC ($(1/R-1)\overline{a}(t,r_t,\theta_t) > 0$). If $\Delta > 0$ fund one exhibits more sensitivity to W_1-risk and will be held long. The second fund is used to neutralize the exposure to W_2-risk induced by the IR-hedge component of fund 1. Combining the IR-hedging demands of the two funds produces a perfect hedge against the impact of IR-risk on the PVC. The overall IR-hedge achieved is positive when risk aversion exceeds 1. This parallels the results found in prior sections.

MPR-hedges display an interesting structure. Since MPRs respond directly to exogenous shocks as well as indirectly through the interest dependence of their drift these MPR-hedges have two components. The direct hedges correspond to the terms with $(\sigma_i^\theta/\sqrt{2\alpha})\overline{b}_i^{dir}(t,r_t,\theta_t)$; indirect hedges involve $\delta_{ir}\overline{b}_i^{ind}(t,r_t,\theta_t)$. Considerations similar to those above govern the signs of these components. Let us focus on the MPR(θ_1)-hedge. We have:

1. direct hedge: fluctuations in the PVC related to the direct impact of (W_1, W_2) on θ_1 are described by the vector

$$(\frac{1}{R}-1)\begin{bmatrix}\alpha\\ \rho_\theta\end{bmatrix}\sigma_1^\theta\frac{1}{\sqrt{2\alpha}}\overline{b}_1^{dir}(t,r_t,\theta_t).$$

When $\Delta > 0$, $\rho_\theta < 0$ and $(1/R-1)\overline{b}_1^{dir}(t,r_t,\theta_t)$ is positive this wealth component is financed by a long (short) position in fund 1 (fund 2). Under these conditions fund 1 is used to span W_1-risk. This, however, will result in an overexposure to W_2-risk. Shorting fund 2 in suitable proportion creates a perfect hedge. When one of the two funds provides a perfect hedge (i.e. $\rho_1/\sqrt{1-\rho_1^2} = \alpha/\rho_\theta$ or $\alpha/\rho_\theta = \rho_2/\sqrt{1-\rho_2^2}$) holdings of the other fund are null.

2. indirect hedge: fluctuations in the PVC induced through the interest rate

dependence of θ_1 are described by

$$(\frac{1}{R}-1)\begin{bmatrix}1\\0\end{bmatrix}\delta_{1r}\overline{b}_1^{ind}(t,r_t,\theta_t)$$

Properties of the fund positions which synthesize this risk parallel the properties of the IR-hedges.

We now provide a numerical illustration of the properties described above as well as others. Consider the symmetric case $(\kappa_1,\overline{\theta}_1,\delta_{1r},\sigma_1) = (\kappa_2,\overline{\theta}_2,\delta_{2r},\sigma_2)$ and calibrate the model by using the parameter estimates reported in section 6: $\kappa_r = 0.06977$, $\overline{r} = 0.005 \times 12$, $\kappa_1 = \kappa_2 = 0.9088$, $\overline{\theta}_1 = \overline{\theta}_2 = 0.1685$, $\delta_{1r} = \delta_{2r} = -23.90/12$, $\sigma_r = 0.01227\sqrt{12}$, $\sigma_1^\theta = \sigma_2^\theta = 0.16127$. The common volatilities of the two funds are $\sigma_1 = \sigma_2 = 0.2$ and the correlation of the market portfolio with the interest rate is $\rho_1 = 0.1$. In order to simulate the portfolio components we set initial values at $r_0 = .06$, $\theta_{10} = \theta_{20} = .10$. The graphs show results for correlations between the MPRs from $\rho_\theta = -0.9$ to $+0.9$ with increment 0.1. Given that $\rho_1 = 0.1$, the correlation ρ_2 of the second risky fund with the interest rate is chosen such that the implied correlation between the risky assets, $\rho = \rho_1\rho_2 + \sqrt{1-\rho_1^2}\sqrt{1-\rho_2^2}$ varies between $-.2$ and $+1$. Values for ρ_2 vary from $\rho_2 = -0.99$ to $+.01$ in increments of 0.01. Finally, risk aversion $R = 4$ and the investment period is taken to be 5 years.

Figures 14 and 15 illustrate, respectively, the behaviors of the mean-variance components and of the IR-hedges. In addition to the theoretical effects described above it should be noted that the IR-hedges are nearly insensitive, and the MV components completely insensitive, to the correlation between the MPRs. Both components increase in magnitude as the returns correlation becomes more positive. Figure 16 shows that the hedge against θ_1 embedded in the demand for fund 1 (fund 2) displays concavity (convexity) with respect to ρ_θ and is increasing (decreasing) with respect to ρ. Figures 17 and 18 reveal that concavity (convexity) reflects the structure of both the direct and indirect components. Furthermore the indirect hedge changes sign when ρ_θ is in a neighborhood of ± 1. This follows from a sign reversal of the covariance $\overline{b}_1^{ind}(t, r_t, \theta_t)$ when ρ_θ approaches ± 1.

The hedge against θ_2 displays surprising behavior (figure 19). Its direct component, in the demand for fund 1, exhibits a convex-concave structure (horizontal S-shape) relative to ρ_θ for large values of ρ; a symmetric pattern characterizes the direct component in the demand for fund 2 (figures 20-21). This S-shape is a consequence of the behavior of the covariance $\overline{b}_2^{dir}(t, r_t, \theta_t)$ which is convex with respect to ρ_θ and takes positive values in a neighborhood of $\rho_\theta = \pm 1$. For values of ρ_θ close to -1 direct hedging entails duplicating the vector $\overline{b}_2^{dir}(t, r_t, \theta_t)(\rho_\theta, \alpha)$ and this is achieved by taking a long position in fund 1. As ρ_θ increases the covariance $\overline{b}_2^{dir}(t, r_t, \theta_t)$ becomes negative which implies a short position in fund 1. As ρ_θ increases further the vector $\overline{b}_2^{dir}(t, r_t, \theta_t)(\rho_\theta, \alpha)$ enters the convex cone formed by asset returns. Both funds are then held long. Eventually, as ρ_θ approaches 1 the covariance $\overline{b}_2^{dir}(t, r_t, \theta_t)$ becomes positive and a short position in fund 1 is required to synthesize $\overline{b}_2^{dir}(t, r_t, \theta_t)(\rho_\theta, \alpha)$.

When combined the total MPR-hedging demand is concave (convex) for fund 1 (fund 2) reflecting the dominance of the hedge against θ_1 (figure 22). Finally figures 23 and 24 show the behavior of the sum of all the hedging components and the overall behavior of the portfolio. The overall hedging demand reflects the reinforcing behaviors of the IR- and MPR-hedges. The overall portfolio structure also exhibits the same pattern. In general hedging implies a significant departure from mean-variance demand behavior.

[Insert figures 14-24 here.]

Numerical values for fund holdings and hedging demands are provided in Table 9 for selected values of the correlation coefficients. They illustrate some of the features discussed above.

[Insert table 9 here.]

2.11 Appendix

2.11.1 Appendix A: proofs

Proof of Theorem 34: It follows from Cox and Huang (1989) and Karatzas, Lehoczky and Shreve (1987) that optimal final wealth must be given by $\hat{X}_T = I(T, \hat{y}\xi_T)$ where $I = [\partial_2 u]^{-1}$ is the inverse marginal utility of consumption and $\hat{y}$ satisfies $\mathbf{E}[\xi_T I(T, \hat{y}\xi_T)] = x$. Since $\xi_t \hat{X}_t = \mathbf{E}_t[\xi_T \hat{X}_T]$ we have for $J(t,y) := yI(t,y)$ that

$$\hat{X}_t = I(t, \hat{y}\xi_t)\mathbf{E}_t[J_{t,T}]$$

where

$$J_{t,T} \equiv \frac{J(T, \hat{y}\xi_T)}{J(t, \hat{y}\xi_t)}.$$

Using the chain rule of Malliavin calculus and the relation $-\partial_2 I(t,y) = \frac{1}{-\partial_{22}u(y, I(t,y))}$ (which follows from the definition $\partial_2 u(t, I(t,y)) = y$) we obtain

$$\frac{\mathcal{D}_s \hat{X}_t}{\hat{X}_t} = -\frac{1}{\xi_t R(t, I(t, \hat{y}\xi_t))}\mathcal{D}_s \xi_t + \frac{\mathcal{D}_s \mathbf{E}_t[J_{t,T}]}{\mathbf{E}_t[J_{t,T}]}$$

where $R(t,x) \equiv \frac{-\partial_{22}u(t,x)x}{\partial_2 u(t,x)}$ is the relative risk aversion of the investor. Taking the limit as $s \uparrow t$ on both sides of this equation and using $\lim_{s\uparrow t} \mathcal{D}_s \hat{X}_t = \hat{\pi}_t' \sigma_t$, $\lim_{s\uparrow t} \mathcal{D}_s \xi_t = -\xi_t \theta_t'$ and the commutativity of the conditional expectation and Malliavin derivative operator then leads to

$$\hat{\pi}_t' = \hat{X}_t \left[\frac{1}{R(t, I(t, \hat{y}\xi_t))}\theta_t' + \frac{\mathbf{E}_t[\mathcal{D}_t J_{t,T}]}{\mathbf{E}_t[J_{t,T}]} \right] \sigma_t^{-1}.$$

But since

$$\mathcal{D}_t J_{t,T} = \frac{\partial_2 J(T,\hat{y}\xi_T)}{J(t,\hat{y}\xi_t)}\mathcal{D}_t\hat{y}\xi_T - \frac{J(T,\hat{y}\xi_T)}{J(t,\hat{y}\xi_t)}\frac{\partial_2 J(t,\hat{y}\xi_t)}{J(t,\hat{y}\xi_t)}\mathcal{D}_t\hat{y}\xi_t$$

where $\mathcal{D}_t\hat{y}\xi_T = -\hat{y}\xi_T(\theta_t' + H_{t,T}')$ with

$$H_{t,T} = \int_t^T \mathcal{D}_t r_s ds + \int_t^T dW_s'\mathcal{D}_t\theta_s + \int_t^T ds\theta_s'\mathcal{D}_t\theta_s = \int_t^T \mathcal{D}_t r_s ds + \int_t^T (dW_s^{\mathbf{Q}})'\mathcal{D}_t\theta_s$$

and since

$$\frac{\partial_2 J(t,y)}{J(t,y)}y = 1 + \frac{y\partial_2 I(t,y)}{I(t,y)} = 1 - \frac{1}{R(t,I(t,y))} := \alpha(t,y)$$

the second term in the expression for the optimal portfolio can be written

$$\begin{aligned}
\frac{\mathbf{E}_t[\mathcal{D}_t J_{t,T}]}{\mathbf{E}_t[J_{t,T}]} &= \frac{1}{\mathbf{E}_t[J_{t,T}]}\mathbf{E}_t\left[\frac{\partial_2 J(T,\hat{y}\xi_T)}{J(t,\hat{y}\xi_t)}\mathcal{D}_t\hat{y}\xi_T - \frac{J(T,\hat{y}\xi_T)}{J(t,\hat{y}\xi_t)}\frac{\partial_2 J(t,\hat{y}\xi_t)}{J(t,\hat{y}\xi_t)}\mathcal{D}_t\hat{y}\xi_t\right] \\
&= -\frac{1}{\mathbf{E}_t[J_{t,T}]}\mathbf{E}_t\left[\frac{\partial_2 J(T,\hat{y}\xi_T)}{J(t,\hat{y}\xi_t)}\hat{y}\xi_T\left(\theta_t' + H_{t,T}'\right)\right] + \frac{\partial_2 J(t,\hat{y}\xi_t)}{J(t,\hat{y}\xi_t)}\hat{y}\xi_t\theta_t' \\
&= -\frac{1}{\mathbf{E}_t[J_{t,T}]}\mathbf{E}_t\left[\frac{J(T,\hat{y}\xi_T)}{J(t,\hat{y}\xi_t)}\alpha(T,I(T,\hat{y}\xi_T))\left(\theta_t' + H_{t,T}'\right)\right] + \alpha(t,I(t,\hat{y}\xi_t))\theta_t' \\
&= \left(\alpha_t - \mathbf{E}_t\left[\frac{J_{t,T}}{\mathbf{E}_t[J_{t,T}]}\alpha_T\right]\right)\theta_t' - \mathbf{E}_t\left[\frac{J_{t,T}}{\mathbf{E}_t[J_{t,T}]}\alpha_T\int_t^T \mathcal{D}_t r_s ds\right] \\
&\quad -\mathbf{E}_t\left[\frac{J_{t,T}}{\mathbf{E}_t[J_{t,T}]}\alpha_T\int_t^T [dW_s + \theta_s ds]'\mathcal{D}_t\theta_s\right]
\end{aligned}$$

where $\alpha_t \equiv \alpha(t,I(t,\hat{y}\xi_t))$ (note that $\alpha_T = \alpha(T,I(T,\hat{y}\xi_T)) = \alpha(T,\widehat{X}_T)$)

Finally using $\frac{J_{t,T}}{\mathbf{E}_t[J_{t,T}]} = \frac{\xi_T}{\xi_t}\frac{\hat{X}_T}{\hat{X}_t}$ where $\frac{d\mathbf{Q}}{d\mathbf{P}}_{|\mathcal{F}_t} = B_t\xi_t$, we obtain

$$\begin{aligned}
\hat{\pi}_t' &= \frac{\hat{X}_t}{R(t,\hat{X}_t)}\mathbf{E}_t^{\mathbf{Q}}\left[\frac{\frac{\hat{X}_T}{B_T}}{\frac{\hat{X}_t}{B_t}}\frac{R(t,\hat{X}_t)}{R(T,\hat{X}_T)}\right]\theta_t'\sigma_t^{-1} \\
&\quad +\hat{X}_t\frac{1-R(t,\hat{X}_t)}{R(t,\hat{X}_t)}\mathbf{E}_t^{\mathbf{Q}}\left[\frac{\frac{\hat{X}_T}{B_T}}{\frac{\hat{X}_t}{B_t}}\frac{R(t,\hat{X}_t)}{R(T,\hat{X}_T)}\left(\frac{R(T,\hat{X}_T)-1}{R(t,\hat{X}_t)-1}\right)\int_t^T \mathcal{D}_t r_s ds\right]\sigma_t^{-1} \\
&\quad +\hat{X}_t\frac{1-R(t,\hat{X}_t)}{R(t,\hat{X}_t)}\mathbf{E}_t^{\mathbf{Q}}\left[\frac{\frac{\hat{X}_T}{B_T}}{\frac{\hat{X}_t}{B_t}}\frac{R(t,\hat{X}_t)}{R(T,\hat{X}_T)}\left(\frac{R(T,\hat{X}_T)-1}{R(t,\hat{X}_t)-1}\right)\int_t^T (dW_s^{\mathbf{Q}})'\mathcal{D}_t\theta_s\right]\sigma_t^{-1}.
\end{aligned}$$

Now note that the chain rule of Malliavin calculus gives

$$\begin{cases} \mathcal{D}_t\theta_s = \partial_2\theta(s,Y_s)\mathcal{D}_t Y_s \\ \mathcal{D}_t r_s = \partial_2 r(s,Y_s)\mathcal{D}_t Y_s \end{cases}.$$

Furthermore (3.2) and Nualart (1995), section 2.2, p. 99-108, imply that $\mathcal{D}_t Y_s = (\mathcal{D}_{1t} Y_s, ..., \mathcal{D}_{dt} Y_s)$ solves d systems (one for each of the d Malliavin derivatives) of d stochastic differential equation

$$\begin{aligned}
\mathcal{D}_{kt} Y_s &= \mathcal{D}_{kt} Y_t + \int_t^s \mathcal{D}_{kt} \mu^Y(v, Y_v) dv + \mathcal{D}_{kt} \int_t^s \left(\sum_{j=1}^d \sigma^Y_{\cdot j}(v, Y_v) dW_{jv} \right) \\
&= \sigma^Y_{\cdot k}(t, Y_t) + \int_t^s \partial_2 \mu^Y(v, Y_v) \mathcal{D}_{kt} Y_v dv + \int_t^s \mathcal{D}_{kt} \left(\sum_{j=1}^d \sigma^Y_{\cdot j}(v, Y_v) dW_{jv} \right) \\
&= \sigma^Y_{\cdot k}(t, Y_t) + \int_t^s \partial_2 \mu^Y(v, Y_v) \mathcal{D}_{kt} Y_v dv + \int_t^s \left(\sum_{j=1}^d \partial_2 \sigma^Y_{\cdot j}(v, Y_v) \mathcal{D}_{kt} Y_v dW_{jv} \right) \\
&= \sigma^Y_{\cdot k}(t, Y_t) + \int_t^s \partial_2 \mu^Y(v, Y_v) \mathcal{D}_{kt} Y_v dv + \int_t^s \left(\sum_{j=1}^d \partial_2 \sigma^Y_{\cdot j}(v, Y_v) dW_{jv} \right) \mathcal{D}_{kt} Y_v
\end{aligned}$$

for $k = 1, ..., d$. The solutions of these systems of linear equations are as stated in the theorem using the fact that the quadratic variation of the martingale part is $\sum_{j=1}^d \partial_2 \sigma^Y_{\cdot j}(v, Y_v)(\partial_2 \sigma^Y_{\cdot j}(v, Y_v))' dv$ where $\sigma^Y_{\cdot\cdot j}$ denotes the j^{th} column of the matrix σ^Y. ■

Proof of Proposition 27: Following the arguments of Doss (1977) we consider a function $F : [0, T] \times \mathbb{R}_+ \mapsto \mathbb{R}_+$ such that $\partial_2 F = \frac{1}{\sigma}$. Using $\partial_{22} F = (\partial_2 \frac{1}{\sigma}) = -\frac{\partial_2 \sigma}{\sigma^2}$ and Ito's lemma implies that

$$dF(t, Y_t) = \left[\frac{\mu}{\sigma} - \frac{1}{2} \partial_2 \sigma + \partial_1 F \right] (t, Y_t) dt + dW_t.$$

so that $F(t, Y_t)$ has the decomposition $F(t, Y_t) = N_t + W_t$ where

$$dN_t = \left[\frac{\mu}{\sigma} - \frac{1}{2} \partial_2 \sigma + \partial_1 F \right] (t, Y_t) dt.$$

Since F has an inverse G given by $G(t, F(t, y)) = y$ we can write $Y_t = G(t, N_t + W_t)$ and therefore

$$dN_t = \left[\frac{\mu}{\sigma} - \frac{1}{2} \partial_2 \sigma + \partial_1 F \right] (t, G(t, N_t + W_t)) dt.$$

with $N_0 = F(0, y)$. Then since from assumptions (i) and (ii) G is continuously differentiable and by theorem 2.2.1 of Nualart (1995) which needs assumption (iii) the process is in the domain of the Malliavin derivative operator $N \in \mathbb{D}^{1,2}$ we have for $t \leq s$ that

$$\mathcal{D}_t Y_s = \partial_2 G(s, N_s + W_s) Z_{t,s}$$

where

$$dZ_{t,s} = \partial_2 \left[\frac{\mu}{\sigma} - \frac{1}{2}\partial_2\sigma + \partial_1 F\right](s, G(s, N_s + W_s))(\partial_2 G(s, N_s + W_s))Z_{t,s} ds$$

with $Z_{t,t} = 1$. Solving this linear SDE for $Z_{t,s}$ and using the relations for derivatives of F and its inverse G produces the result stated. ■.

Proof of Proposition 28: Since $dY_t = \mu(t, Y_t)dt + \sigma(t, Y_t)dW_t$ we have

$$\int_0^t \theta(s, Y_s) dW_s = -\int_0^t [\frac{\theta}{\sigma}\mu](s, Y_s)ds + \int_0^t [\frac{\theta}{\sigma}](s, Y_s)dY_s.$$

Then for ψ such that $\partial_2\psi\sigma = \theta$ we have that

$$\psi(t, Y_t) - \psi(0, Y_0) = \int_0^t [\partial_1\psi + \frac{1}{2}\partial_{22}\psi\sigma^2](s, Y_s)ds + \int_0^t [\frac{\theta}{\sigma}](s, Y_s)dY_s.$$

But $\partial_{22}\psi = \frac{\partial_2\theta}{\sigma} - \frac{\theta}{\sigma}\frac{\partial_2\sigma}{\sigma}$ and therefore

$$\int_0^t \theta(s, Y_s)dW_s = -\int_0^t [\frac{\theta}{\sigma}\mu](s, Y_s)ds + \psi(t, Y_t) - \psi(0, Y_0) - \int_0^t [\partial_1\psi + \frac{1}{2}[\partial_2\theta\sigma - \theta\partial_2\sigma]](s, Y_s)d$$

Using this expression for the stochastic integral in the expression of the SPD provides (2.18).

To establish (2.19) use $\int_t^T \mathcal{D}_t\theta_s[dW_s + \theta_s ds] = \mathcal{D}_t\{\int_0^t \theta_s[dW_s + \frac{1}{2}\theta_s ds]\}$, substitute the expression for $\int_0^t \theta(s, Y_s)dW_s$ above on the right hand side, and compute the Malliavin derivative of the expression in bracket. ■

Proof of Corollary 30: Substituting $\gamma_r, \gamma_\theta = 0$ in the expressions for the Malliavin derivatives in Proposition 27 gives $\mathbf{D}_t r_v = \sigma_r \exp[-\kappa_r(v - t)]$ and

$\mathbf{D}_t\theta_v = \sigma_\theta \exp[-\kappa_\theta(v-t)]$. Since R is constant and $\mathbf{D}_t r_v$ is deterministic we can then write

$$a(t,r_t,S_t) = E_t\left[\frac{\xi_{t,T}^{1-1/R}}{E_t\left[\xi_{t,T}^{1-1/R}\right]}\left(\int_t^T \sigma_r \exp\left[-\kappa_r(v-t)\right]dv\right)\right] = \int_t^T \sigma_r \exp\left[-\kappa_r(v-t)\right]dv.$$

Substituting the expression for $\mathbf{D}_t\theta_v$ in $b(t,r_t,\theta_t)$ gives the formula in the lemma. ■

Proof of equations (2.42)-(2.44): We conjecture that the individual price of W_2-risk is null. The SPD is then given by the formula in theorem 26 where (r,θ) satisfy (2.39)-(2.40). Since (r,θ) is independent of W_2-risk, optimal wealth $\widehat{X}_T = I(T,\widehat{y}\xi_T)$ is independent of W_2. The Martingale representation theorem and the Clark-Ocone formula imply the existence of a unique financing portfolio which is given by (2.42)-(2.44). ■

Proof of equation (2.49): Theorem 26 implies that the optimal portfolio is given by

$$\hat{\pi}_t = \hat{X}_t(\sigma_t')^{-1}\left[\frac{1}{R}\theta_t + (\frac{1}{R}-1)a(t,r_t,\theta_t) + (\frac{1}{R}-1)b_1(t,r_t,\theta_t) + (\frac{1}{R}-1)b_2(t,r_t,\theta_t)\right]$$

where

$$a(t,r_t,\theta_t)' \equiv \mathbf{E}_t\left[\frac{\xi_{t,T}^{1-1/R}}{\mathbf{E}_t\left[\xi_{t,T}^{1-1/R}\right]}\int_t^T \mathcal{D}_t r_s ds\right]$$

$$b_i(t,r_t,\theta_t)' \equiv \mathbf{E}_t\left[\frac{\xi_{t,T}^{1-1/R}}{\mathbf{E}_t\left[\xi_{t,T}^{1-1/R}\right]}\int_t^T \mathcal{D}_t\theta_{is} dW_{is}^{\mathbf{Q}}\right];\ i=1,2.$$

Straightforward computations give the Malliavin derivatives

$$(\mathcal{D}_t r_s)' = \begin{bmatrix}\mathcal{D}_{1t}r_s \\ \mathcal{D}_{2t}r_s\end{bmatrix} = \begin{bmatrix}-\sigma_r r_s^{\frac{1}{2}}\exp\left(-\frac{1}{2}\int_t^s(\kappa_r - \frac{\sigma_r^2}{4})\frac{1}{r_s}ds - \frac{1}{2}\kappa_r(s-t)\right) \\ 0\end{bmatrix}$$

$$(\mathcal{D}_t\theta_{1s})' = \begin{bmatrix} \mathcal{D}_{1t}\theta_{1s} \\ \mathcal{D}_{2t}\theta_{1s} \end{bmatrix} = \begin{bmatrix} \sigma_1^\theta \frac{\alpha}{\sqrt{2\alpha}} e^{-\kappa_1(s-t)} + \int_t^s e^{-\kappa_1(s-v)}\delta_{1r}\mathcal{D}_{1t}r_v dv \\ \sigma_1^\theta \frac{\rho_\theta}{\sqrt{2\alpha}} e^{-\kappa_1(s-t)} \end{bmatrix}$$

$$(\mathcal{D}_t\theta_{2s})' = \begin{bmatrix} \mathcal{D}_{1t}\theta_{2s} \\ \mathcal{D}_{2t}\theta_{2s} \end{bmatrix} = \begin{bmatrix} \sigma_2^\theta \frac{\rho_\theta}{\sqrt{2\alpha}} e^{-\kappa_2(s-t)} + \int_t^s e^{-\kappa_2(s-v)}\delta_{2r}\mathcal{D}_{1t}r_v dv \\ \sigma_1^\theta \frac{\alpha}{\sqrt{2\alpha}} \end{bmatrix}$$

which leads to

$$a(t, r_t, \theta_t) = \begin{bmatrix} \mathbf{E}_t\left[\frac{\xi_{t,T}^{1-1/R}}{\mathbf{E}_t\left[\xi_{t,T}^{1-1/R}\right]}\int_t^T \mathcal{D}_{1t}r_s ds\right] \\ \mathbf{E}_t\left[\frac{\xi_{t,T}^{1-1/R}}{\mathbf{E}_t\left[\xi_{t,T}^{1-1/R}\right]}\int_t^T \mathcal{D}_{2t}r_s ds\right] \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \mathbf{E}_t\left[\frac{\xi_{t,T}^{1-1/R}}{\mathbf{E}_t\left[\xi_{t,T}^{1-1/R}\right]}\int_t^T \mathcal{D}_{1t}r_s ds\right]$$

$$\begin{aligned} b_1(t, r_t, \theta_t) &= \begin{bmatrix} \alpha \\ \rho_\theta \end{bmatrix} \sigma_1^\theta \frac{1}{\sqrt{2\alpha}} \mathbf{E}_t\left[\frac{\xi_{t,T}^{1-1/R}}{\mathbf{E}_t\left[\xi_{t,T}^{1-1/R}\right]}\int_t^T e^{-\kappa_1(s-t)} dW_{1s}^{\mathbf{Q}}\right] \\ &+ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \delta_{1r}\mathbf{E}_t\left[\frac{\xi_{t,T}^{1-1/R}}{\mathbf{E}_t\left[\xi_{t,T}^{1-1/R}\right]}\int_t^T \int_t^s e^{-\kappa_1(s-v)}\mathcal{D}_{1t}r_v dv dW_{1s}^{\mathbf{Q}}\right] \end{aligned}$$

and a symmetric expression for $b_2(t, r_t, \theta_t)$.

Defining the determinant of the volatility matrix $\Delta = \sigma_1\sigma_2\left(\rho_1\sqrt{1-\rho_2^2} - \rho_2\sqrt{1-\rho_1^2}\right)$ we can write the mean-variance term as

$$\frac{1}{R}(\sigma_t')^{-1}\theta_t = \frac{1}{R}\frac{1}{\Delta}\begin{bmatrix} \sigma_2\sqrt{1-\rho_2^2} & -\sigma_2\rho_2 \\ -\sigma_1\sqrt{1-\rho_1^2} & \sigma_1\rho_1 \end{bmatrix}\begin{bmatrix} \theta_{1t} \\ \theta_{2t} \end{bmatrix}.$$

Substituting the expression for $a(t, r_t, \theta_t)$ gives the IR-hedge

$$(\frac{1}{R}-1)(\sigma_t')^{-1}a(t, r_t, \theta_t) = (\frac{1}{R}-1)\frac{1}{\Delta}\begin{bmatrix} \sigma_2\sqrt{1-\rho_2^2} \\ -\sigma_1\sqrt{1-\rho_1^2} \end{bmatrix}\mathbf{E}_t\left[\frac{\xi_{t,T}^{1-1/R}}{\mathbf{E}_t\left[\xi_{t,T}^{1-1/R}\right]}\int_t^T \mathcal{D}_{1t}r_s ds\right].$$

Finally, substituting $b_1(t, r_t, \theta_t)$ provides the MPR(θ_1)-hedge

$$\begin{aligned} &(\frac{1}{R}-1)(\sigma_t')^{-1}b_1(t, r_t, \theta_t) \\ = \; &(\frac{1}{R}-1)\frac{1}{\Delta}\begin{bmatrix} \sigma_2\sqrt{1-\rho_2^2}\alpha - \rho_\theta\sigma_2\rho_2 \\ -\sigma_1\sqrt{1-\rho_1^2}\alpha + \rho_\theta\sigma_1\rho_1 \end{bmatrix}\sigma_1^\theta\frac{1}{\sqrt{2\alpha}}\mathbf{E}_t\left[\frac{\xi_{t,T}^{1-1/R}}{\mathbf{E}_t\left[\xi_{t,T}^{1-1/R}\right]}\int_t^T e^{-\kappa_1(s-t)}dW_{1s}^{\mathbf{Q}}\right] \\ &+(\frac{1}{R}-1)\frac{1}{\Delta}\begin{bmatrix} \sigma_2\sqrt{1-\rho_2^2} \\ -\sigma_1\sqrt{1-\rho_1^2} \end{bmatrix}\delta_{1r}\mathbf{E}_t\left[\frac{\xi_{t,T}^{1-1/R}}{\mathbf{E}_t\left[\xi_{t,T}^{1-1/R}\right]}\int_t^T\int_t^s e^{-\kappa_1(s-v)}\mathcal{D}_{1t}r_v dv dW_{1s}^{\mathbf{Q}}\right]. \end{aligned}$$

A symmetric expression holds for the MPR(θ_2)-hedge. ■

2.11.2 Appendix B: A representation of Malliavin derivatives of multivariate diffusion processes

Consider a d-dimensional process Y which satisfies the system of SDEs

$$dY_t = \mu(t, Y_t)dt + \sigma(t, Y_t)dW_t; Y_0 = y$$

where W is a d-dimensional Brownian motion process. For any $d \times 1$ vector of functions $f(t, Y)$ let $\partial_1 f$ represent the $d \times 1$ vector of first derivatives relative to time and $\partial_2 f$ the $d \times n$ matrix whose rows are composed of the gradients relative to Y of the elements of f. The Malliavin derivative of Y has the following alternative representation.

Proposition 31 *If the following conditions hold*

(i) differentiability of drift: $\mu \in \mathcal{C}^1([0,T] \times \mathbb{R}^d)$

(ii) differentiability of volatility: $\sigma \in \mathcal{C}^2([0,T] \times \mathbb{R}^d)$

(iii) growth condition: $\mu(t, 0)$ and $\sigma(t, 0)$ are bounded for all $t \in [0, T]$

(iv) invertibility condition: $det(\sigma(t, y)) \neq 0$ for all $t \in [0,T]$ and $y \in \mathbb{R}^d$

(v) volatility condition: the Lie algebra of the vector fields generated by the columns of σ, $\mathcal{L}\{\sigma_1, \dots, \sigma_d\}$ is Abelian, i.e. $(\partial_2\sigma_i)\sigma_j = (\partial_2\sigma_j)\sigma_i$ for all $i, j = 1, \dots, d$ where $\partial_2\sigma_j$ is the $d \times d$ Jacobian matrix with respect to y of the $d \times 1$ vector function σ_j.

then we have for $t \leq s$ that

$$\mathcal{D}_t Y_s = \sigma(s, Y_s) Z_{t,s}$$

where the $d \times d$ process $Z_{t,s}$ satisfies

$$dZ_{t,s} = \left[\partial_2 \left[(\sigma)^{-1}\mu + \frac{1}{2}H\right] + \partial_1(\sigma)^{-1}\right](s, Y_s)\sigma(s, Y_s)Z_{t,s}ds \tag{2.50}$$

subject to the boundary condition $Z_{t,t} = I_d$ ($d \times d$-identity matrix) where

$$H = (I \otimes \mathbf{1}')(K \odot (\mathbf{1} \otimes \sigma'\sigma))\mathbf{1} \tag{2.51}$$

with K for the Jacobian matrix of σ^{-1} given by

$$K = -\frac{1}{2}[(\sigma \otimes \sigma')^{-1}\partial_2(\sigma') + [(\partial_2\sigma')'(\sigma' \otimes \sigma)^{-1}]_\nu]. \tag{2.52}$$

The operators $\otimes$ and $\odot$ represent, respectively, the Kronecker and Hadamard products,[23] whereas the stack operator $[\cdot]_\nu$ operates on a $d \times d^2$ matrix $B = [B_1, \ldots, B_d]$ where B_i are d-dimensional square matrices as follows: $[B]_\nu = [(B_1)', \ldots, (B_d)']'$.

Assumption (v) in this proposition guarantees that there exists F such that $\partial_2 F = \sigma^{-1}$. Since by (iv) F has an inverse G, say, condition (v) could equivalently be written as $\partial_j G_i(t, z) = \sigma_{i,j}(t, G(t, z))$. The assumption is automatically satisfied if the state variables do not interact with each other, i.e. if $\sigma_j(t, Y_t) = \sigma_j(t, Y_t^j)$ for $j = 1, ..., d$. The one dimensional case treated earlier falls in this category.

Proof of Proposition 31: The proof parallels the one dimensional case. Assumption (v) ensures the existence of a $d \times 1$ vector of functions $F : [0, T] \times \mathbb{R}_+^d \mapsto \mathbb{R}_+^d$ such that $\partial_2 F = \sigma^{-1}$. Using $\partial_{22} F = \partial_2 \sigma^{-1}$ we get by the identification theorem for Hessian matrices of vector functions (theorem 6.7. of Magnus and Neudecker (1988)) that

$$\partial_{22}F(t, y) = -\frac{1}{2}[(\sigma \otimes \sigma')^{-1}\partial_2(\sigma') + [(\partial_2\sigma')'(\sigma' \otimes \sigma)^{-1}]_\nu](t, y) \tag{2.53}$$

[23]The Kronecker product of a vector Y and a matrix $A = [a_{ij}]$ is $X \otimes A = [Ya_{ij}]$. The Hadamard product of two matrices A and B is $A \odot B = [a_{i,j}b_{i,j}]$, i.e. the matrix composed of the direct products of the corresponding elements in the two matrices.

where the stack operator $[\cdot]_\nu$ acts in the following manner: for a $d \times d^2$ matrix $B = [B_1, \dots, B_d]$ where B_i are d-dimensional square matrices we have $[B]_\nu = [(B_1)', \dots, (B_d)']'$. The use of the stack operator is necessary to guarantee that the components $\partial_{22}F_i(t,y)$ which arise in blocks in $\partial_{22}F(t,y)$ remain symmetric.

Using Ito's lemma applied to each element of F we get

$$dF_i(t,Y_t) = [\partial_1 F_i + \partial_2 F_i \mu + \frac{1}{2} trace(\partial_{22}F_i \sigma' \sigma)](t,Y_t)dt + [\partial_2 F_i \sigma](t,Y_t)dW_t$$

for $i = 1, ..., d$. Stacking these SDEs for $i = 1, \dots, d$ one below the other gives for $F(t,Y_t)$,

$$dF(t,Y_t) = \left[\sigma^{-1}\mu + \frac{1}{2}H + \partial_1 F\right](t,Y_t)dt + dW_t$$

where $H' = [trace(\partial_{22}F_1\sigma'\sigma), \dots, trace(\partial_{22}F_d\sigma'\sigma)]$. To obtain the expression (2.51) for H note that $trace(AB') = \mathbf{1}'(A \odot B)\mathbf{1}$ where $\odot$ is the Hadamard product, i.e. $A \odot B = [[a_{i,j}b_{i,j}]]$. Now we can write the matrix H as follows

$$H = [\mathbf{1}'((\partial_{22}F_1)' \odot \sigma'\sigma)\mathbf{1}, \dots, \mathbf{1}'((\partial_{22}F_d)' \odot \sigma'\sigma)\mathbf{1}]'$$

which is equivalent to

$$H = [(I \otimes \mathbf{1}')[((\partial_{22}F_1)' \odot \sigma'\sigma), \dots, ((\partial_{22}F_d)' \odot \sigma'\sigma)]'\mathbf{1}.$$

But since $\partial_{22}F = [(\partial_{22}F_1)', \dots, (\partial_{22}F_d)']'$ we get

$$[((\partial_{22}F_1)' \odot \sigma'\sigma), \dots, ((\partial_{22}F_d)' \odot \sigma'\sigma)]' = (\partial_{22}F \odot (\mathbf{1} \otimes \sigma'\sigma))$$

and therefore $H = (I \otimes \mathbf{1}')(\partial_{22}F \odot (\mathbf{1} \otimes \sigma'\sigma))\mathbf{1}$ where $\partial_{22}F(t,y)$ is as given in (2.53). Thus, H is obtained by multiplying the Hessian of each element of F element by element with the matrix $\sigma'\sigma$ then summing over all elements and arranging the result in a column vector whose first element is obtained by performing the same operation for F_1, the second for F_2 and so on until F_d. This establishes (2.51).

Thus, using these expressions we see that $F(t, Y_t) = N_t + W_t$ where

$$dN_t = \left[\sigma^{-1}\mu + \frac{1}{2}H + \partial_1 F\right](t, Y_t)dt.$$

Since the determinant of the Jacobian $\partial_2 F$ differs from 0 (assumption (iv)) the vector $F(t, y)$ has a unique inverse G defined by $G(t, F(t, y)) = y$. We can then write $Y_t = G(t, N_t + W_t)$ and therefore

$$dN_t = \left[\sigma^{-1}\mu + \frac{1}{2}H + \partial_1 F\right](t, G(t, N_t + W_t))dt$$

with $N_0 = F(0, y)$. Then since from assumptions (i)-(ii) G is continuously differentiable and by theorem 2.2.1 of Nualart (1995), which requires assumption (iii), the process is in the domain of the Malliavin derivative operator $N \in \mathbb{D}^{1,2}$ we have for $t \leq s$ that

$$\mathcal{D}_t Y_s = \partial_2 G(s, N_s + W_s) Z_{t,s}$$

where

$$dZ_{t,s} = \partial_2 \left[\sigma^{-1}\mu + \frac{1}{2}H + \partial_1 F\right](s, G(s, N_s + W_s))(\partial_2 G(s, N_s + W_s))Z_{t,s}ds$$

with $Z_{t,t} = I_d$. Since $\partial_2 G(s, N_s + W_s)\partial_2 F(s, Y_s) = I$ we have that $\partial_2 G(s, N_s + W_s) = \sigma(s, Y_s)$. Substituting in the equation above leads to the result in the proposition. ■

2.11.3 Appendix C: the MRGID model

We consider the following interest rate - market price of risk model with interaction in the drift of the MPR

$$dr_t = \kappa_r(\overline{r} - r_t)dt + \sigma_r\sqrt{r_t}dW_t, \qquad r_0 \text{ given} \tag{2.54}$$

$$d\theta_t = \left(\kappa_\theta(\overline{\theta} - \theta_t) + \delta_\theta r_t\right)dt + \sigma_\theta dW_t, \qquad \theta_0 \text{ given} \tag{2.55}$$

where $(\kappa_r, \overline{r}, \sigma_r, \kappa_\theta, \bar{\theta}, \delta_\theta, \sigma_\theta)$ are nonnegative constants. The transition from the general model with state variables Y to the model (2.54)-(2.55) with state variables (r, θ) is immediate since the Malliavin derivative $\mathbf{D}_t\theta_v$ can now be computed directly from the process (2.55). Taking account of the specific structure (2.54)-(2.55) then leads to

Proposition 32 *In the financial market (2.54)-(2.55) the optimal portfolio is given by (2.5) with*

$$\mathbf{D}_t r_v = \sqrt{r}_v \sigma_r \exp\left[-\frac{1}{2}\int_t^v \left(\kappa_r(1+\overline{r}\frac{1}{r_u}) - \frac{1}{4}\sigma_r^2(\frac{1}{r_u})\right) du\right]$$

$$\mathbf{D}_t\theta_v = \sigma_\theta e^{-\kappa_\theta(v-t)} + \delta_\theta \int_t^v e^{-\kappa_\theta(v-s)}\mathcal{D}_t r_s ds.$$

The SPD is then

$$\xi_t = \exp\left[-\int_0^t [r_s + (\frac{1}{2} - \frac{\kappa_\theta}{\sigma_\theta})\theta_s^2 - \frac{\kappa_\theta}{\sigma_\theta}\bar{\theta}\theta_s - \frac{\delta}{\sigma_\theta}\theta_s r_s]ds] - \frac{1}{2}(\theta_t^2 - \theta_0^2) + \frac{1}{2}\sigma_\theta t\right]$$

and the stochastic integral for the MPR-hedge becomes

$$\int_t^T \mathcal{D}_t\theta_s[dW_s + \theta_s ds] = \int_t^T \left[[(1+2\frac{\kappa_\theta}{\sigma_\theta}) - \frac{\kappa_\theta}{\sigma_\theta}\bar{\theta} + \frac{\delta}{\sigma_\theta}r_s]\mathcal{D}_t\theta_s + \frac{\delta}{\sigma_\theta}\theta_s\mathcal{D}_t r_s\right] ds + \frac{1}{\sigma_\theta}\theta_T\mathcal{D}_t\theta_T - \theta_t.$$

2.11.4 Appendix D: asymptotic laws of state variables estimators

In this appendix we report theorems from Detemple, Garcia and Rindisbacher (2000) providing the asymptotic laws of estimators of functionals of Brownian motions. The proofs of these results are based on Kurtz and Protter (1991) and Jacod and Protter (1998). Consider the SDE of the vector of state variables Y_t after the Doss transformation

$$d\widehat{Y}_t = \widehat{m}(\widehat{Y}_t)dt + \sum_{j=1}^{d} dW_t^j \tag{2.56}$$

with

$$\hat{m}(Y_s) = [(\sigma^Y)^{-1}]_i'\mu^Y + \frac{1}{2}tr[\partial[(\sigma^Y)^{-1}]_i'\sigma^Y(\sigma^Y)'](Y_s), \tag{2.57}$$

and let $\widehat{Y}_T^N$ denote the estimator of $\widehat{Y}_T$ based on a Euler scheme. Our next theorem characterizes the estimation error.

Theorem 33 *The asymptotic law of the estimator of the state variables Y is given by*

$$U_{t,T}^{\widehat{Y}^N} \equiv N(\widehat{Y}_T^N - Y_T) \Rightarrow \hat{U}_{t,T}^Y$$

where

$$\hat{U}_{t,T}^Y = -\sigma^Y(Y_T)\hat{\Omega}_{t,T}\int_t^T \hat{\Omega}_{t,s}^{-1}\partial\hat{m}(Y_s)\left[\frac{1}{2}(\hat{m}(Y_s)ds + dW_s) + \frac{1}{\sqrt{12}}dB_s\right]. \tag{2.58}$$

with

$$\hat{\Omega}_{t,v} = \exp\left(\int_t^v [(\partial\hat{m})\sigma^Y](Y_s)ds\right). \tag{2.59}$$

In addition to providing an explicit expression for the asymptotic law of the estimator, theorem 40 also demonstrates a speed of convergence of order $1/N$. These results can be contrasted with those obtained when state variables are estimated before transformation. Applying a Euler scheme to estimate the solution of (3.2) leads to

$$U_t^{\widetilde{Y}^N} \equiv \sqrt{N}(\widetilde{Y}_t^N - Y_t) \Rightarrow \tilde{U}_t^Y$$

where

$$\tilde{U}_{t,v}^Y = -\frac{1}{\sqrt{2}}\tilde{\Omega}_{t,v}\int_t^v \tilde{\Omega}_{t,s}^{-1}\sum_{h,j=1}^d [(\partial\sigma_{\cdot j}^Y)\sigma_{\cdot h}^Y](Y_s)dB_s^{h,j}. \tag{2.60}$$

with

$$\tilde{\Omega}_{t,v} = \exp\left(\int_t^v [\partial\mu^Y(Y_s) - \frac{1}{2}\sum_{j=1}^d(\partial\sigma_{\cdot j}^Y(Y_s))^2]ds + \sum_{j=1}^d\int_t^v \partial\sigma_{\cdot j}^Y(Y_s)dW_s^j\right). \tag{2.61}$$

In this case the resulting speed of convergence is $1/\sqrt{N}$. These results illustrate the increase in the speed of convergence achieved by using the Doss transformation. They also highlights the fact that the limit law is different and involves an exponential of a bounded total variation process instead of a stochastic integral. DGR (2000) provides similar theorems for the Malliavin derivatives and the functionals that appear in the hedging terms $a(t, Y_t)$ and $b(t, Y_t)$. The increased rate of convergence is important when computing conditional estimators of the hedging terms based on an approximation of the dynamic evolution of the state variables.

2.11.5 Appenix E: Tables

Table 1- Comparison of the speeds of convergence of the discretization schemes when the IR follows a MRSR process.

N	r		$\mathcal{D}r$	
	Euler	Euler-Transform	Euler	Euler-Transform
2	0.000115598	5.49255e-06	5.81463e-07	3.47457e-07
4	0.000111128	3.37985e-06	3.58341e-07	2.13681e-07
8	8.74541e-05	1.82631e-06	2.33208e-07	1.15422e-07
16	6.50156e-05	9.41716e-07	1.6312e-07	5.9616e-08
32	4.66084e-05	4.7979e-07	1.16983e-07	3.03396e-08
64	3.336e-05	2.40698e-07	8.29213e-08	1.52396e-08
128	2.3761e-05	1.20386e-07	5.97503e-08	7.63041e-09
256	1.68824e-05	5.83759e-08	4.18739e-08	3.69586e-09
512	1.19618e-05	2.53747e-08	3.00371e-08	1.60477e-09

Table 2 - Unconstrained monthly estimates of the bivariate interest rate-MPR process with constant stock volatility

Parameters	ML estimates	Standard Errors
κ_{rM}	0.0265	0.0107
$\overline{r}$	0.0053	0.0007
κ_{θ}	0.6528	0.0482
$\overline{\theta}$	0.0846	0.0084
σ_r	0.0049	0.0002
σ_{θ}	0.1052	0.0039
$\rho_{r\theta}$	-0.1651	0.0539

Table 3 - Constrained (with $\rho_{r\theta}$ set at -0.9) monthly estimates of the bivariate interest rate-MPR process with constant stock volatility

Parameters	ML estimates	Standard Errors
κ_{rM}	0.0824	0.0116
$\overline{r}$	0.0050	0.0005
κ_{θ}	0.6950	0.0507
$\overline{\theta}$	0.0871	0.0161
σ_r	0.0105	0.0004
σ_{θ}	0.2125	0.0080

Table 4 - Constrained (with $\rho_{r\theta}$ set at -0.9) monthly estimates of the bivariate interest rate-MPR process with constant stock volatility with r_{t-1} in the drift of MPR

Parameters	ML estimates	Standard Errors
κ_{rM}	0.0005	0.0185
$\bar{r}$	0.0051	0.0010
κ_θ	0.7771	0.0484
$\bar{\theta}$	0.2675	0.0348
σ_r	0.0105	0.0004
σ_θ	0.2050	0.0073
δ	-26.2469	4.9686

Table 5 - Unconstrained monthly estimates of the bivariate interest rate-MPR process with a GARCH stock conditional variance

Parameters	ML estimates	Standard Errors
κ_{rM}	0.0290	0.0106
$\bar{r}$	0.0053	0.0006
κ_θ	0.5975	0.0464
$\bar{\theta}$	0.0882	0.0083
σ_r	0.0049	0.0002
σ_θ	0.0979	0.0035
$\rho_{r\theta}$	-0.1863	0.052

Table 6 - Constrained (with $\rho_{r\theta}$ set at -0.9) monthly estimates of the bivariate interest rate-MPR process with a GARCH

stock conditional variance

Parameters	ML estimates	Standard Errors
κ_{rM}	0.0947	0.0128
$\overline{r}$	0.0050	0.0004
κ_θ	0.6826	0.0507
$\overline{\theta}$	0.0900	0.0147
σ_r	0.0104	0.0004
σ_θ	0.1928	0.0070

Table 7 - Shares of the portfolio in the stock and Hedging Components for Model 1.

$R=2$	Investment horizon	1	2	3	4	5
	Stock demand	25.4	26.1	27.0	29.2	30.5
	MPR-hedge	-1.7	-3.0	-3.9	-3.5	-3.7
	Interest rate hedge	2.1	4.1	5.9	7.6	9.2

$T=1$	Risk aversion	0.5	1	1.5	4	5
	Stock demand	113.0	50.0	33.2	14.4	12.3
	MPR-hedge	17.2	0.0	-1.6	-1.3	-1.24
	Interest rate hedge	-4.3	0.0	1.4	3.2	3.4

Table 8 - Dividend-Price Ratio Model - Shares of the portfolio in the stock and Hedging Components for Model 1 ($\sigma = 0.20$).

$R=4$	Investment horizon	1	2	3	4	5
	Stock demand	30.18	39.92	46.88	52.45	57.27
	MPR-hedge	13.90	20.13	23.82	26.35	28.33
	Interest rate hedge	3.78	7.29	10.56	13.60	16.43

$T=1$	Risk aversion	0.5	1	1.5	2	5
	Stock demand	81.42	50.0	40.86	36.49	28.94
	MPR-hedge	-13.56	0.0	5.85	8.98	14.91
	Interest rate hedge	-5.02	0.0	1.68	2.52	4.03

Table 9 - Multiasset model - Shares invested

in the two Funds and Hedging Components (case $\rho_2 < 0$).

Returns Correlation: $\rho = 0$							
Fund	ρ_θ	MV-Comp.	IR-H	MPR1-H	MPR2-H	H-Comp.	Holdings
1	-0.9	13.68594	1.62469	-0.46143	-2.86238	-1.69912	11.98682
2		-11.18734	-16.18569	-0.54965	-2.24462	-18.97997	-30.16731
1	-0.6	13.68594	1.61552	-1.85189	3.16808	2.93171	16.61764
2		-11.18734	-16.09436	-8.47951	1.43846	-23.13540	-34.32275
1	-0.3	13.68594	1.61200	-0.56593	6.75339	7.79946	21.48540
2		-11.18734	-16.05926	-12.29448	1.77988	-26.57386	-37.76121
1	0.0	13.68594	1.61197	1.28385	6.38567	9.28149	22.96742
2		-11.18734	-16.05898	-12.79016	0.67745	-28.17169	-39.35903
1	0.3	13.68594	1.61506	2.55858	4.96426	9.13790	22.82384
2		-11.18734	-16.08978	-10.21850	-0.23006	-26.53834	-37.72568
1	0.6	13.68594	1.61690	3.73219	2.63805	7.98715	21.67309
2		-11.18734	-16.10812	-8.57007	-0.57762	-25.25582	-36.44316
1	0.9	13.68594	1.62484	1.22769	-3.01071	-0.15819	13.52775
2		-11.18734	-16.18717	-1.46995	1.47530	-16.18182	-27.36916

Returns Correlation: $\rho = .5$							
Fund	ρ_θ	MV-Comp.	IR-H	MPR1-H	MPR2-H	H-Comp.	Holdings
1	-0.9	20.14629	10.97144	-0.14402	-1.56618	9.26124	29.40754
2		-12.91800	-18.68958	-0.63468	-2.59186	-21.91612	-34.83412
1	-0.6	20.14629	10.90953	3.04477	2.33741	16.29171	36.43800
2		-12.91800	-18.58412	-9.79127	1.66099	-26.71440	-39.63239
1	-0.3	20.14629	10.88574	6.53376	5.72556	23.14507	43.29136
2		-12.91800	-18.54359	-14.19641	2.05522	-30.68478	-43.60277
1	0.0	20.14629	10.88555	8.66978	5.99446	25.54979	45.69608
2		-12.91800	-18.54326	-14.76876	0.78225	-32.52978	-45.44777
1	0.3	20.14629	10.90643	8.45946	5.09711	24.46300	44.60929
2		-12.91800	-18.57883	-11.79928	-0.26565	-30.64375	-43.56175
1	0.6	20.14629	10.91886	8.68116	2.97161	22.57163	42.71792
2		-12.91800	-18.60001	-9.89584	-0.66698	-29.16283	-42.08083
1	0.9	20.14629	10.97244	2.07654	-3.86265	9.18633	29.33262
2		-12.91800	-18.69129	-1.69734	1.70352	-18.68511	-31.60311

Returns Correlation: $\rho = .9$							
Fund	ρ_θ	MV-Comp.	IR-H	MPR1-H	MPR2-H	H-Comp.	Holdings
1	-0.9	36.78663	35.04645	0.67354	1.77253	37.49252	74.27915
2		-25.66581	-37.13294	-1.26099	-5.14958	-43.54350	-69.20931
1	-0.6	36.78663	34.84868	15.65741	0.19780	50.70389	87.49052
2		-25.66581	-36.92339	-19.45354	3.30010	-53.07684	-78.74264
1	-0.3	36.78663	34.77270	24.82088	3.07813	62.67170	99.45833
2		-25.66581	-36.84288	-28.20578	4.08337	-60.96529	-86.6311
1	0.0	36.78663	34.77207	27.69419	4.98680	67.45306	104.23969
2		-25.66581	-36.84222	-29.34296	1.55420	-64.63098	-90.2967
1	0.3	36.78663	34.83877	23.65871	5.43931	63.93679	100.7234
2		-25.66581	-36.91289	-23.44310	-0.52779	-60.88378	-86.5495
1	0.6	36.78663	34.87849	21.42850	3.83078	60.13777	96.9244
2		-25.66581	-36.95497	-19.66131	-1.32517	-57.94146	-83.6072
1	0.9	36.78663	35.04965	4.26297	-6.05705	33.25558	70.0422
2		-25.66581	-37.13633	-3.37232	3.38461	-37.12405	-62.78985

2.11.6 Appendix F: Graphics

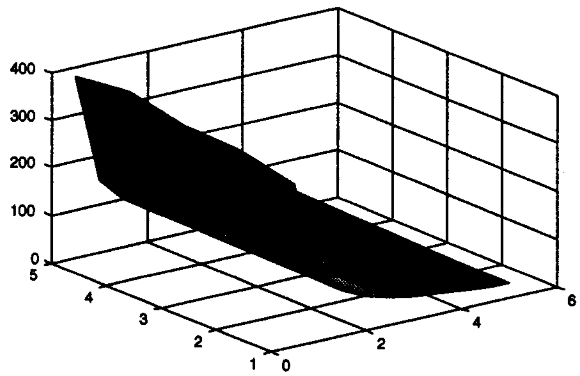

Figure 1: Share of portfolio invested in stock as a function of time and risk aversion.

R=2	Investment horizon	1	2	3	4	5
	Stock demand	72.7	73.2	74.4	76.7	78.3

T=1	Risk aversion	0.5	1	3	4	5
	Stock demand	339.1	150.0	48.9	37.3	30.5

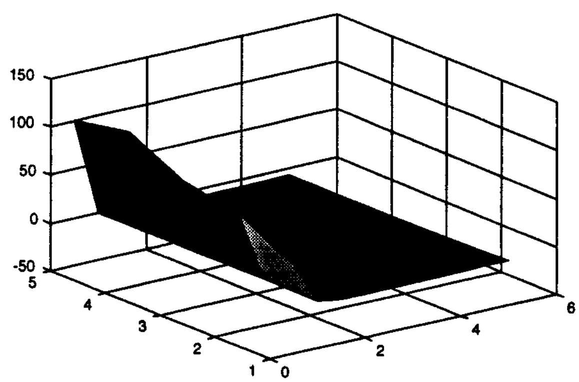

Figure 2: Share of the MPR hedge as a function of time and risk aversion.

R=2	Investment horizon	1	2	3	4	5
	MPR-hedge	-4.5	-6.3	-7.4	-7.3	-8

T=1	Risk aversion	.05	1	3	4	5
	MPR-hedge	43.5	0.0	−4.1	−3.5	-3.5

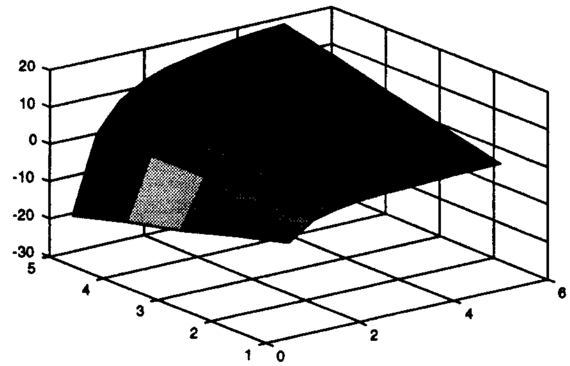

Figure 3: Share of the interest rate hedge as a function of time and risk aversion.

R=2	Investment horizon	1	2	3	4	5
	Interest rate hedge	2.2	4.5	6.7	9.0	11.3

T=1	Risk aversion	0.5	1	3	4	5
	Interest rate hedge	-4.4	0.0	3.0	3.4	3.6

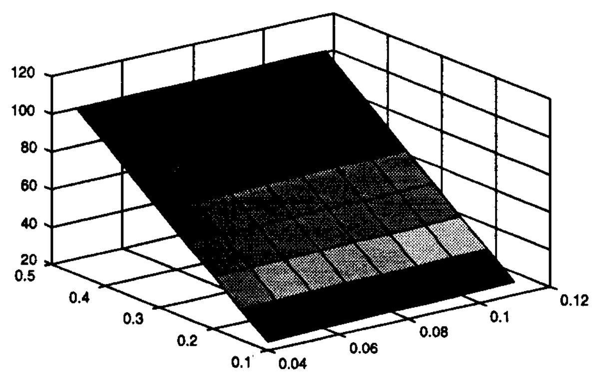

Figure 4: Stock demand behavior relative to r_0 and θ_0. Interest rate varies between 0.04 and 0.08; MPR between .05 and .40.

r=6%	MPR	0.10	0.20	0.40
	Stock demand	25.4	49.1	96.4

MPR=.25	Interest rate (%)	4	6	8
	Stock demand	60.3	60.9	61.5

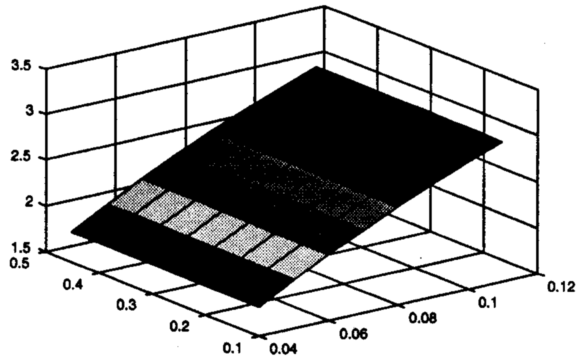

Figure 5: Interest rate hedge behavior relative to r_0 and θ_0. Interest rate varies between 0.04 and 0.08; MPR between .05 and .40.

r=6%	MPR	0.10	0.20	0.40
	Interest rate hedge	2.2	2.2	2.2

MPR=.25	Interest rate(%)	4	6	8
	Interest rate hedge	1.8	2.2	2.6

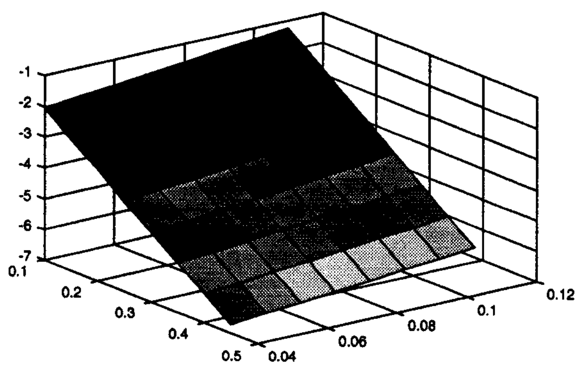

Figure 6: MPR -hedge behavior relative to r_0 and θ_0. Interest rate varies between 0.04 and 0.11; MPR between .10 and .45.

r=6%	MPR	0.10	0.20	0.40
	MPR-hedge	-1.8	-3.2	-5.8

MPR=.25	IR(%)	4	6	8
	MPR-hedge	-4.0	-3.8	-3.6

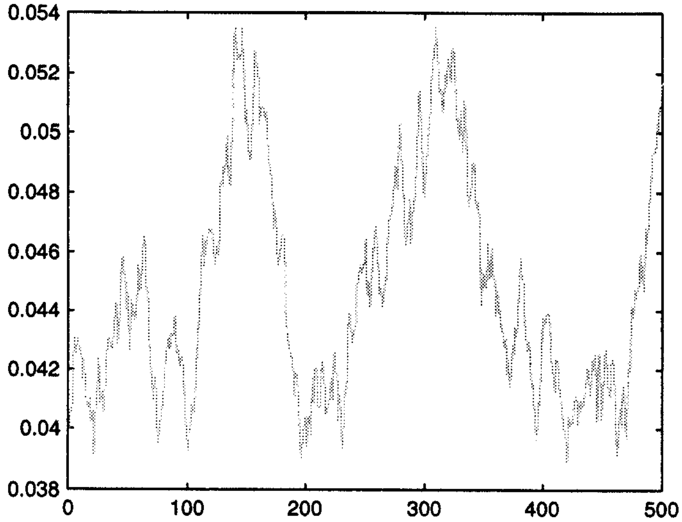

Figure 7: Simulated Path for Interest Rate.

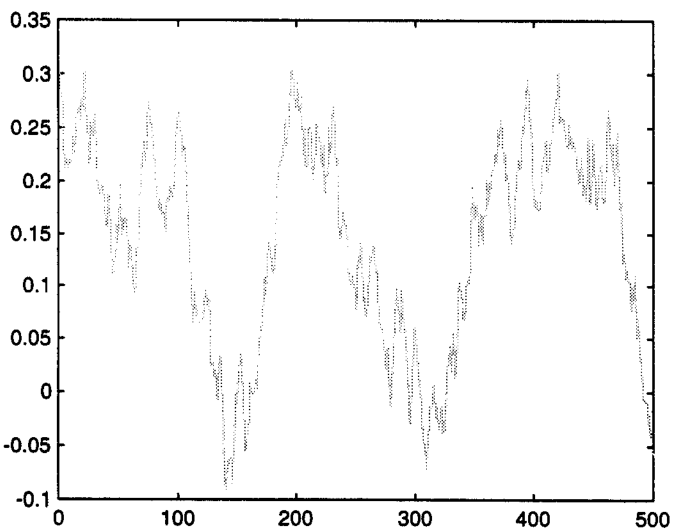

Figure 8: Simulated Path for Market Price of Risk.

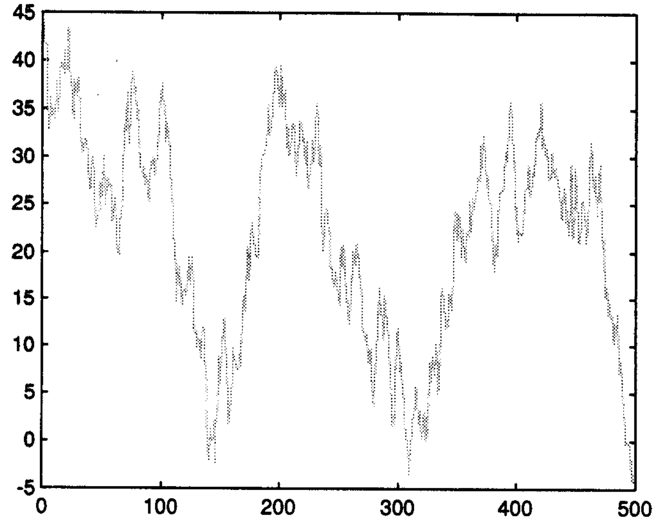

Figure 9: Fixed 5-year Horizon - Share in Stocks - R = 4.

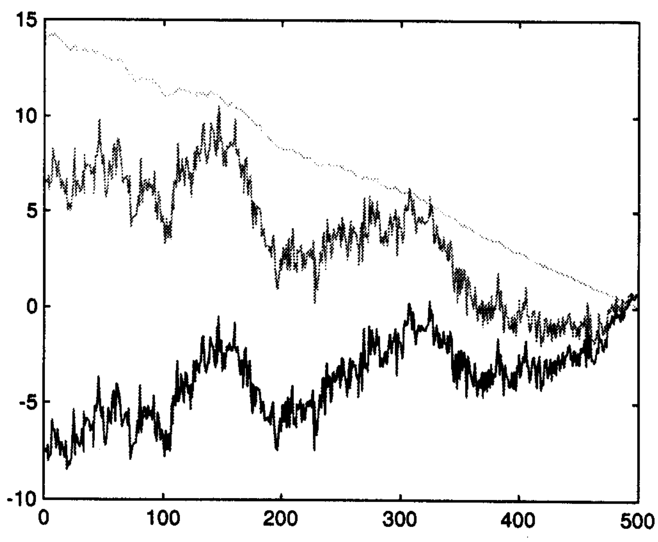

Figure 10: Fixed 5-year Horizon - Hedging Shares (top to bottom): Interest Rate, Total, MPR-R = 4.

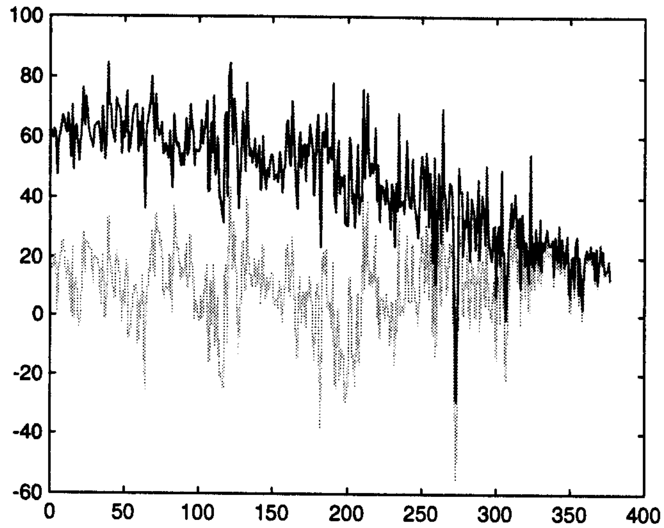

Figure 11: Share of Stock in Portfolio with (top) and without (bottom) hedging - Fixed Horizon of 31.5 years (our sample).

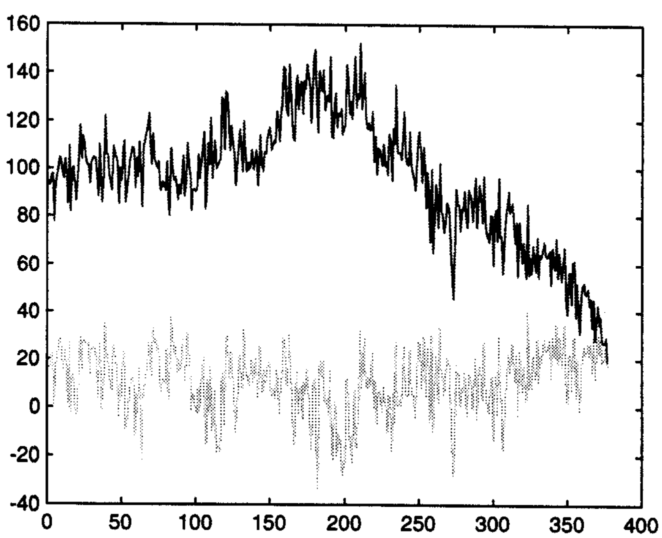

Figure 12: Model with Stochastic Dividends - Share of Stock in Portfolio with (top) and without (bottom) hedging - Fixed Horizon of 31.5 years (our sample).

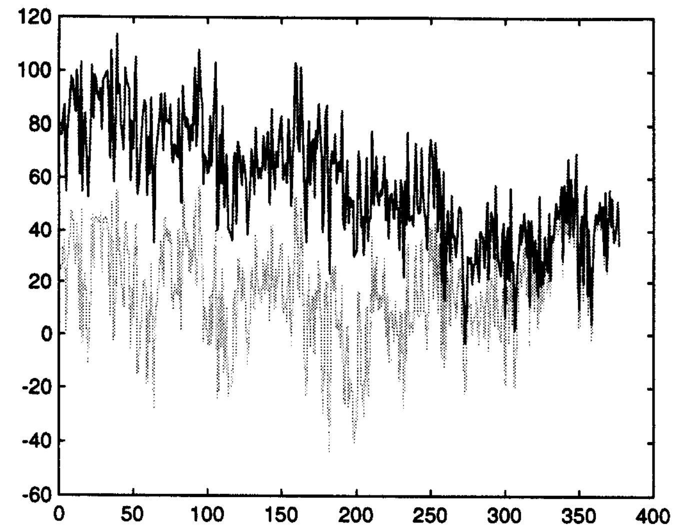

Figure 13: Model with Stochastic Volatility - Share of Stock in Portfolio with (top) and without (bottom) hedging - Fixed Horizon of 31.5 years (our sample).

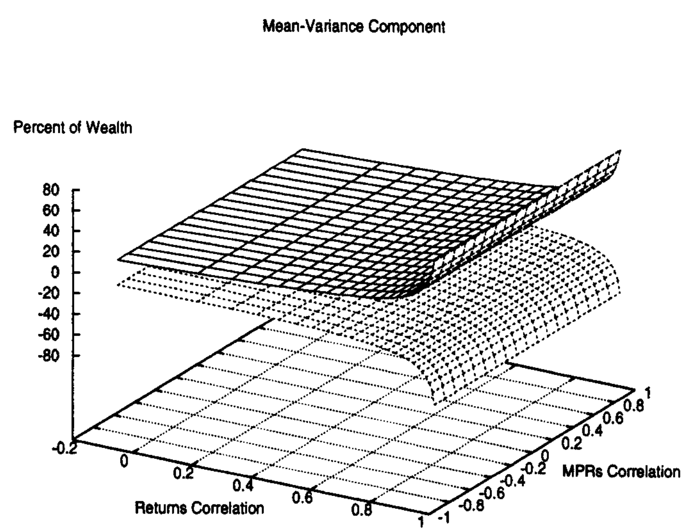

Figure 14: Mean-Variance Component: Fund 1 (plain) and Fund 2 (dotted line)

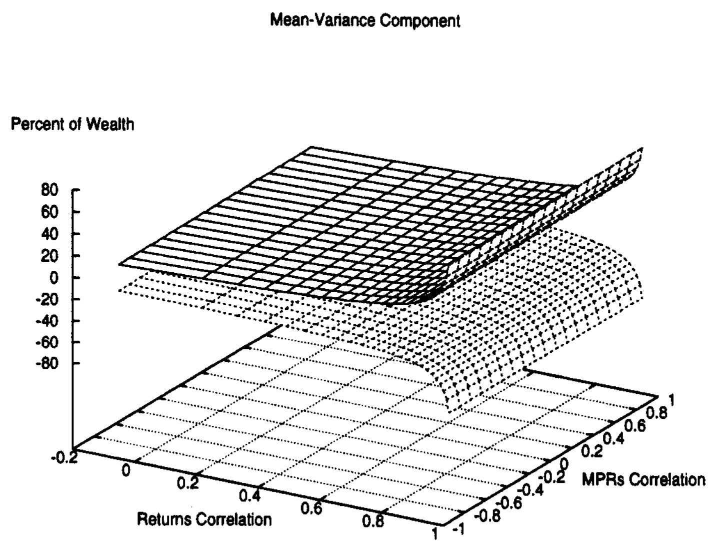

Figure 2.1: Mean-Variance Component: Fund 1 (plain) and Fund 2 (dotted line)

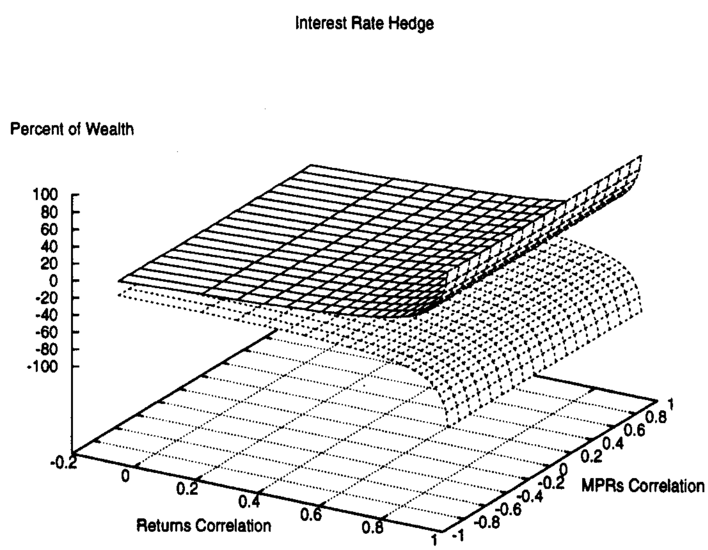

Figure 2.2: Interest Rate Hedge: Fund 1 (plain) and Fund 2 (dotted line)

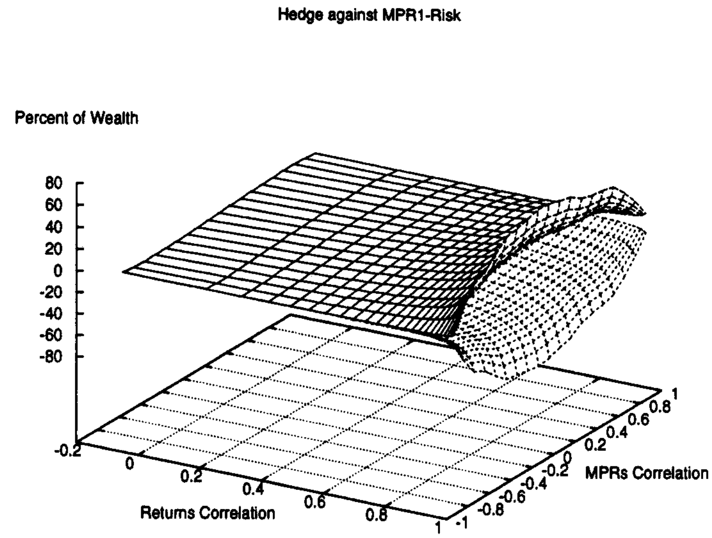

Figure 2.3: Hedge against MPR1-Risk: Fund 1 (plain) and Fund 2 (dotted line)

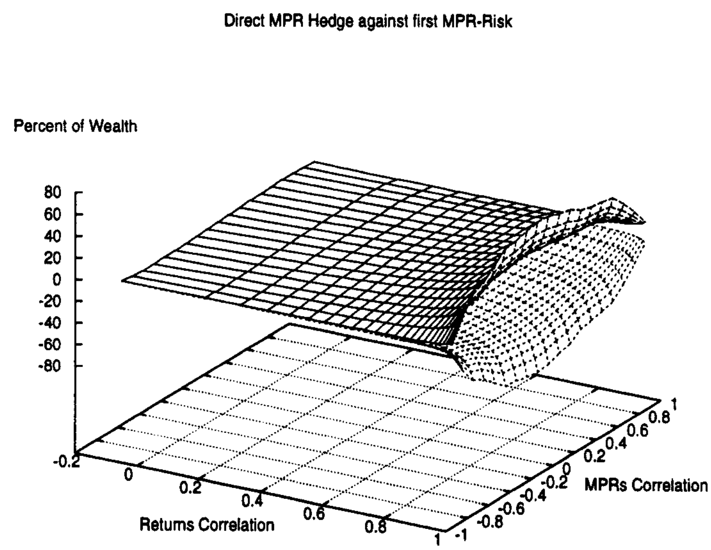

Figure 2.4: Direct MPR Hedge against first MPR-Risk: Fund 1 (plain) and Fund 2 (dotted line)

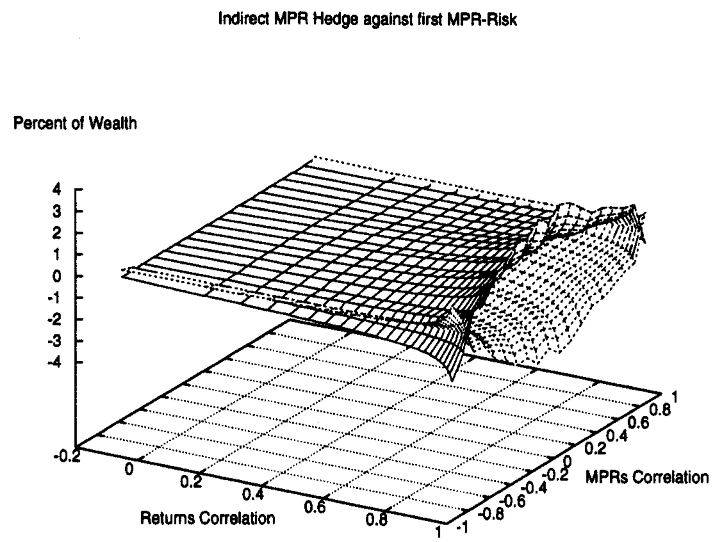

Figure 2.5: Indirect MPR Hedge against first MPR-Risk: Fund 1 (plain) and Fund 2 (dotted line)

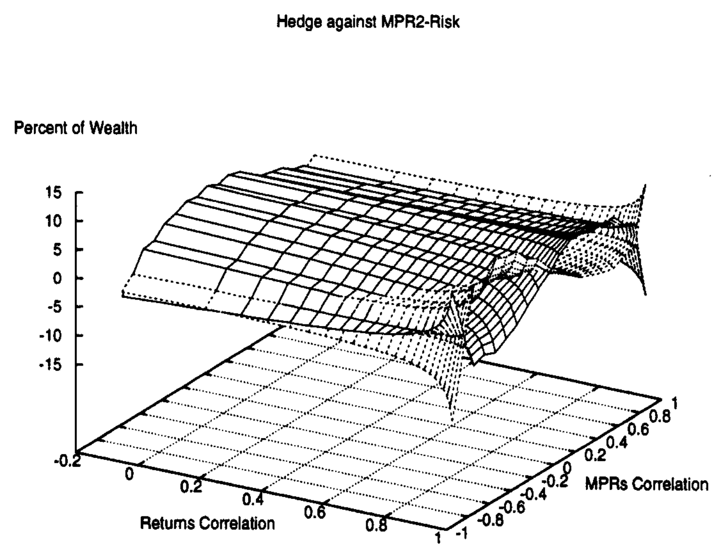

Figure 2.6: Hedge against MPR2-Risk: Fund 1 (plain) and Fund 2 (dotted line)

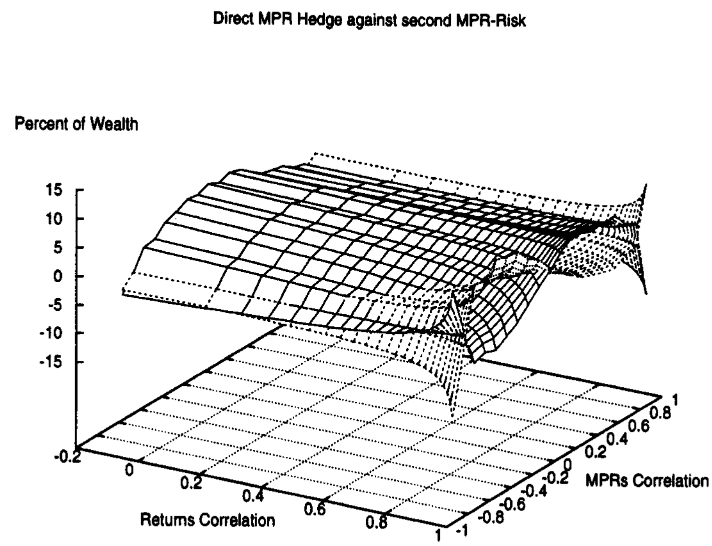

Figure 2.7: Direct MPR Hedge against second MPR-Risk: Fund 1 (plain) and Fund 2 (dotted line)

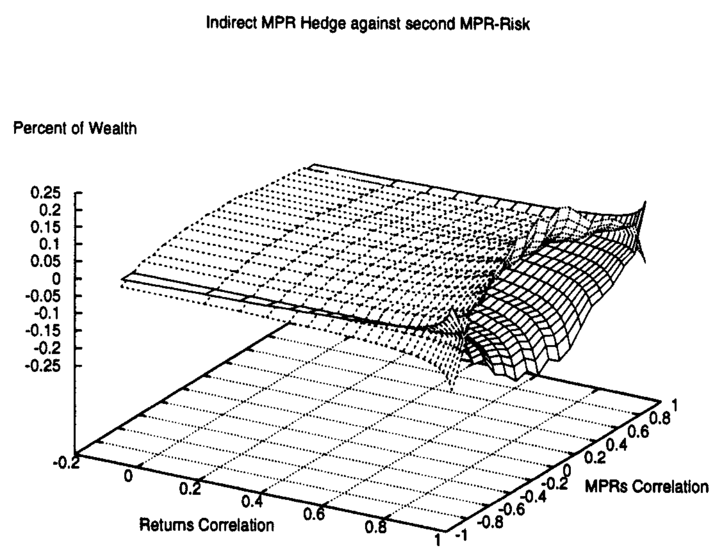

Figure 2.8: Indirect MPR Hedge against second MPR-Risk: Fund 1 (plain) and Fund 2 (dotted line)

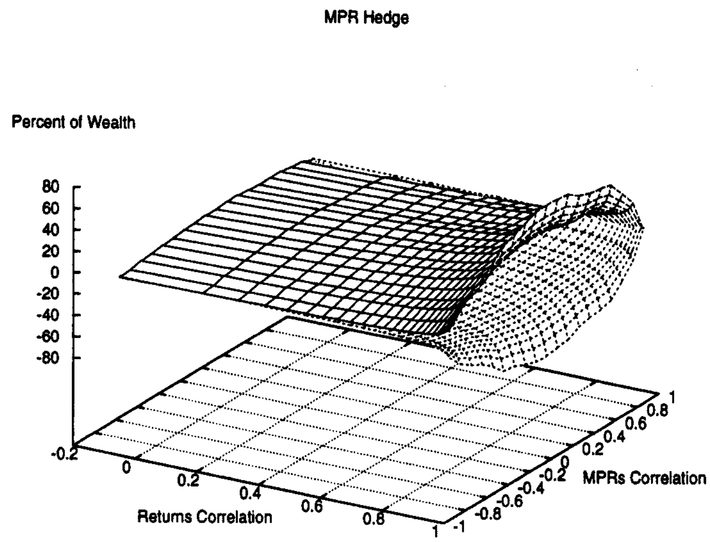

Figure 2.9: MPR Hedge: Fund 1 (plain) and Fund 2 (dotted line)

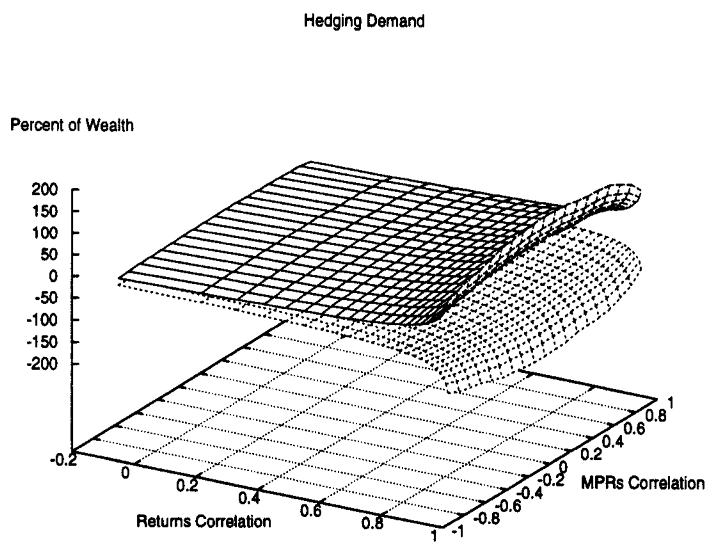

Figure 2.10: Hedge Demand: Fund 1 (plain) and Fund 2 (dotted line)

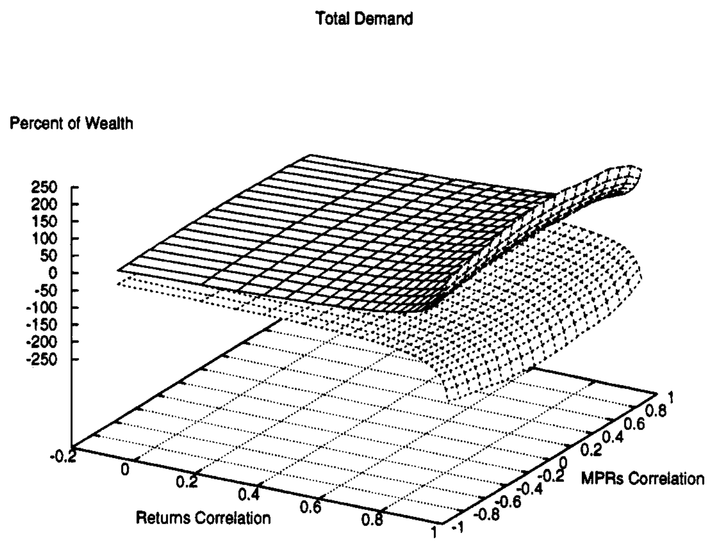

Figure 2.11: Total Demand: Fund 1 (plain) and Fund 2 (dotted line)

Chapter 3

Asymptotic Properties of Optimal Portfolio Estimators

3.1 Optimal Portfolios

The motivation for our study stems from results on optimal portfolios. To set the stage we review key results introducing the random processes to be estimated.

Consider a financial market in which uncertainty is captured by d Brownian motions $W_t^i, i = 1, ..., d$. Two types of assets are traded: d risky stocks and 1 riskless asset. The stock prices $S_i, i = 1, ..., d$, satisfy the stochastic differential equations

$$dS_{it} = S_{it}[(\mu_i(t, Y_t) - \delta_i(t, Y_t))dt + \sigma_i(t, Y_t)dW_t]; \quad i = 1, ..., d \tag{3.1}$$

$$dY_t = \mu^Y(t, Y_t)dt + \sigma^Y(t, Y_t)dW_t \tag{3.2}$$

where Y is a d-dimensional vector of state variables describing the evolution of the opportunity set. Here μ_i is the gross expected return, δ_i the dividend rate and σ_i the vector of volatility coefficients of security i. The riskless asset pays an interest rate $r_t = r(t, Y_t)$. We assume that the coefficients $r(t, Y_t)$, $\mu_i(t, Y_t)$, $\delta_i(t, Y_t)$, $\mu^Y(t, Y_t)$, respectively, $\sigma^Y(t, Y_t)$ are integrable $(P - a.s.)$, that $\sigma_i(t, Y_t)$ is square-integrable $(P - a.s.)$ and that all the coefficients satisfy Growth and Lipschitz conditions for the existence of a weak solution to (3.1)-(3.2). Let σ denote the $d \times d$-dimensional volatility matrix whose rows are $\sigma_i, i = 1, ..., d$. Assume that σ is nonsingular almost everywhere and that the market price of risk process

$$\theta_t = \theta(t, Y_t) = \sigma(t, Y_t)^{-1}(\mu(t, Y_t) - r(t, Y_t)\mathbf{1}),$$

where $\mathbf{1}$ is the unit vector, is continuously differentiable and satisfies the Novikov condition $E \exp\left(\frac{1}{2}\int_0^T \theta_t'\theta_t dt\right) < \infty$. Under this condition the risk neutral measure is given by $dQ = \eta_T dP$ where

$$\eta_t = \exp\left[-\int_0^t \theta_t' dW_t - \frac{1}{2}\int_0^t \theta_t'\theta_t dt\right].$$

The state price density is $\xi_t \equiv B_t^{-1}\eta_t$ where $B_t \equiv \exp[\int_0^t r_s ds]$ is the date t-value of a dollar investment in the riskless asset. Under Q the process $W_t^{\mathbf{Q}} = W_t + \int_0^t \theta_v dv$ is a Brownian motion. Let $\xi_{t,v} \equiv \xi_v/\xi_t$ denote the relative state price density.

The portfolio choice problem of an investor consists in selecting a dynamic portfolio policy so as to maximize the expected utility of terminal wealth

$$\max_{\pi} E\left[\frac{1}{1-R}X_T^{1-R}\right] \quad s.t. \tag{3.3}$$

$$\begin{cases} dX_t = r_t X_t dt + X_t \pi_t'[(\mu_t - r_t\mathbf{1})dt + \sigma_t dW_t], \qquad X_0 = x \\ X_t \geq 0 \text{ for all } t \in [0, T]. \end{cases} \tag{3.4}$$

Here X_t represents wealth at date t, x is initial wealth and π_t the vector of proportions invested in the risky assets at t. The nonnegativity constraint prevents bankruptcy. In (3.3) we focus immediately on a specialized utility function with constant relative risk aversion $R > 0$.

This portfolio choice problem was originally analyzed by Merton (1969, 1971) using dynamic programming techniques. Karatzas, Lehoczky and Shreve (1987) and Cox-Huang (1989)) have proposed a resolution procedure based on martingale methods which applies to a larger class of (Ito) price processes. In this context, Ocone and Karatzas (1991) have used the Clark-Ocone formula to write the optimal portfolio in terms of abstract Malliavin derivatives of the coefficients of the model. A specialization to diffusion processes, which identifies the Malliavin derivatives in explicit form, is found in Detemple, Garcia and Rindisbacher (2000). For constant relative risk aversion they show that the optimal portfolio is

Theorem 34 *Let $\rho = 1 - 1/R$ and suppose that $E(\xi_T^\rho) < \infty$ and $\xi_T^\rho \in \mathbb{D}^{1,2}$. Then*

$$\hat{\pi}_t = (\sigma(t, Y_t)')^{-1} \left[\frac{1}{R}\theta(t, Y_t) + \left(\frac{1}{R} - 1\right) a(t, Y_t) + \left(\frac{1}{R} - 1\right) b(t, Y_t)\right] \tag{3.5}$$

where

$$a(t, Y_t)' \equiv \mathbf{E}_t \left[\frac{\xi_{t,T}^\rho}{\mathbf{E}_t\left[\xi_{t,T}^\rho\right]} \int_t^T \mathcal{D}_t r_s ds\right] \tag{3.6}$$

$$b(t, Y_t)' \equiv \mathbf{E}_t \left[\frac{\xi_{t,T}^\rho}{\mathbf{E}_t\left[\xi_{t,T}^\rho\right]} \int_t^T (dW_s^{\mathbf{Q}})' \mathcal{D}_t \theta_s\right]. \tag{3.7}$$

The Malliavin derivatives in (3.6)-(3.7) are $\mathcal{D}_t\theta_s = \partial_2\theta(s, Y_s)\mathcal{D}_t Y_s$ and $\mathcal{D}_t r_s = \partial_2 r(s, Y_s)\mathcal{D}_t Y_s$ where

$$\mathcal{D}_t Y_s = \sigma^Y(t, Y_t) \exp\left(\int_t^s dL_v\right), \tag{3.8}$$

with the $k \times k$ random variable dL_v defined by[1]

$$dL_v \equiv \left[\partial_2 \mu^Y(v, Y_v) - \frac{1}{2}\sum_{j=1}^{d} \partial_2 \sigma^Y_{\cdot j}(v, Y_v)(\partial_2 \sigma^Y_{\cdot j}(v, Y_v))'\right] dv + \sum_{j=1}^{d} \partial_2 \sigma^Y_{\cdot j}(v, Y_v) dW^j_v$$

and $\sigma^Y_{\cdot j}$ denotes the j^{th} column of σ^Y and $\partial_2 f(v, Y) \equiv \partial f(v, Y)/\partial Y$.

Expression (3.5) shows that the optimal portfolio has 3 components: a mean-variance term ($\theta(t, Y_t)$), a hedging term against interest rate fluctuations ($a(t, Y_t)$) and a hedging term against market price of risk fluctuations ($b(t, Y_t)$). Since the last two terms involve expectations of functionals of the Malliavin derivatives of the state variables Y we need numerical methods that approximate such objects.

An alternative representation, in the spirit of Doss (1977), can be obtained by performing a change of (state) variables,

Theorem 35 *(Detemple, Garcia and Rindisbacher (2000)). Suppose that there exists functions $G : \times\mathbb{R}^d \to \mathbb{R}^d$ and $\Phi : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$ such that*

(i) $\partial_{z_j} G_i(t, z) = \sigma_{ij}(t, G(z))$ for $i, j = \{1, ..., d\}$

(ii) $\partial_{z_j} \Phi(t, z) = \theta_j(t, G(z))$ for $j = \{1, ..., d\}$,

and let F be the inverse of G, i.e. $F(t, G(t, z)) = z$.. The new state variables $Z_{it} \equiv F_i(t, Y_t)$, $i = 1, ..., d$ satisfy the SDEs

$$Z_{it} = F_i(0, x) + \int_0^t \hat{\mu}_i(s, Z_s) ds + W^i_t \tag{3.9}$$

where $\hat{\mu}_i = \hat{m}_i \circ G$ with

$$\hat{m}_i \equiv [(\sigma^Y)^{-1}]'_i(\mu^Y - \frac{1}{2}\sum_{j=1}^{d}(\partial \sigma^Y_{\cdot j})\sigma^Y_{\cdot j}) + \partial_1 F_i \tag{3.10}$$

[1] The exponential in (3.8) should be interpreted as the exponential of a matrix, i.e. (3.8) is short hand notation for the solution of $d\mathcal{D}_t Y_s = \left(dL_s + \frac{1}{2} d[L]_s\right) \mathcal{D}_t Y_s$ subject to the boundary condition $\mathcal{D}_t Y_t = \sigma^Y(t, Y_t)$, where $[L]$ is the quadratic variation process.

and where $[(\sigma^Y)^{-1}]_i'$ *is the* i^{th} *row of the inverse of* σ^Y*. The optimal portfolio is given by (3.5)-(3.7) with*

$$\mathcal{D}_t Y_s = \sigma(s, Y_s) \exp\left[\int_t^s [\partial_2 \mu \sigma^Y - (\sigma^Y)^{-1} \partial_1 \sigma^Y](s, Y_s)\right] \tag{3.11}$$

Theorem 35 provides an alternative representation of the random variables arising in the hedging components. This formula is obtained by passing to a new system of state variables, Z, whose evolution has unit volatility. The advantage of this transformation (hereafter called Doss transformation) is at the computational level. Unit volatility implies the existence of an exact Euler discretization scheme which avoids biases associated with the martingale part of the SDE and improves the speed of convergence in simulation exercises.

While the benefits of Doss transformations in terms of speed of convergence for the solutions of SDEs is straightforward, little is known about convergence properties of functionals of those solutions, such as those arising in the hedging terms. The purpose of the next sections is to study these properties and to establish the limit laws of these functionals and of the corresponding estimators for the hedging terms.

The results presented are based on techniques developped recently by Jacod and Protter (1998) and Kurtz and Protter (1999). Without loss of generality we will restrict the analysis to the case of homogeneous dynamics of state variables (3.2).

3.2 Embedding of the Problem

In this section we perform an expansion of the set of state variables in order to write the random variables and functionals entering the functions a and b as the solutions of systems of SDEs. To this end define the functions,

Definition 1: *For* $x' = [x_1, \dots, x_6] \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}$ *let*

$$h(x_3) \equiv [r + \tfrac{1}{2}\|\theta\|^2](x_3)$$

$$\hat{m}_i(x_3) = [[(\sigma^Y)^{-1}]_i'(\mu^Y - \tfrac{1}{2}\textstyle\sum_{j=1}^d (\partial\sigma_{\cdot j}^Y)\sigma_j^Y)](x_3)$$
$$i = 1, \dots, d$$

$$\hat{h}(x_3) \equiv [\theta'\hat{m} + \tfrac{1}{2}tr[(\partial\theta)\sigma^Y]](x_3)$$

where $[(\sigma^Y)^{-1}]_i'$ *is the ith row of the inverse of* σ^Y.

The k^{th} $(k = 1, \dots, d)$ element of the a and b entering the hedging terms can be rewritten in the form

$$a_k(t, Y_t)' \equiv \mathbf{E}_t \left[\frac{\xi_{t,T}^\rho}{\mathbf{E}_t\left[\xi_{t,T}^\rho\right]} \int_t^T \mathcal{D}_{k,t} r_s ds \right] = \mathbf{E}_t \left[\frac{H_{2,k,t,T}}{\mathbf{E}_t\left[H_{1,t,T}\right]} \right]$$

$$b_k(t, Y_t)' \equiv \mathbf{E}_t \left[\frac{\xi_{t,T}^\rho}{\mathbf{E}_t\left[\xi_{t,T}^\rho\right]} \int_t^T (dW_s^{\mathbf{Q}})' \mathcal{D}_{k,t}\theta_s \right] = \mathbf{E}_t \left[\frac{H_{3,k,t,T}}{\mathbf{E}_t\left[H_{1,t,T}\right]} \right]$$

where $H_{1,t,T} = J_{0,t,T}J_{1,t,T}$, $H_{2,k,t,T} = H_{1,t,T}J_{2,k,t,T}$ respectively $H_{3,k,t,T} = H_{1,t,T}J_{3,k,t,T}$ with

$$J_{0,t,v} \equiv \exp\left(-\rho \textstyle\int_t^v h(Y_s)ds\right)$$

$$J_{1,t,v} \equiv \exp\left(-\textstyle\int_t^v [\rho\theta(Y_s)'dW_s + \tfrac{1}{2}\|\rho\theta(Y_s)\|^2 ds]\right)$$

$$J_{2,k,t,v} \equiv \textstyle\int_t^v \partial r(Y_s)\mathcal{D}_{k,t}Y_s ds \qquad k = 1, \dots, d$$

$$J_{3,k,t,v} \equiv \textstyle\int_t^v \theta(Y_s)'\mathcal{D}_{k,t}Y_s ds + \rho\int_t^v (dW_s)'\partial\theta(Y_s)\mathcal{D}_{k,t}Y_s \qquad k = 1, \dots, d.$$

By Ito's lemma the vector

$$\tilde{X}_{k,t,T}' \equiv [J_{0,t,v}, J_{1,t,v}, Y_t', (\mathcal{D}_{k,t}Y_t)', J_{2,k,t,v}, J_{3,k,t,v}]$$

satisfies the integral equation

$$\tilde{X}_{k,t,T} = \tilde{X}_{k,t,t} + \int_t^T \tilde{A}(\tilde{X}_{k,t,s})ds + \sum_{j=1}^d \int_t^T \tilde{B}_j(\tilde{X}_{k,t,s})dW_s^j \qquad (3.12)$$

with coefficients $\tilde{A} \in \mathcal{C}^1(\mathbb{R}^{2d+4})$ and $\tilde{B}_j \in \mathcal{C}^1(\mathbb{R}^{2d+4})$ given by

$$\tilde{A}(x) = \begin{bmatrix} -x_1 h(x_3) \\ -\frac{1}{2}x_2\|\rho\theta(x_3)\|^2 \\ \mu^Y(x_3) \\ \partial\mu^Y(x_3)x_4 \\ \rho\partial r(x_3)x_4 \\ -[\partial h - \rho\partial r](x_3)x_4 \end{bmatrix} \quad \text{and} \quad \tilde{B}(x) = \begin{bmatrix} 0 \\ -\rho x_2\theta_j(x_3) \\ \sigma_{\cdot j}^Y(x_3) \\ \partial\sigma_{\cdot j}^Y(x_3)x_4 \\ 0 \\ \partial\theta_j(x_3)x_4 \end{bmatrix}$$

and initial condition $\tilde{X}_{k,t,t} = [1, 1, Y_t', \sigma_{\cdot k}^Y(Y_t)', 0, 0]$

In a similar manner we can also rewrite the expressions of theorem 35 as

$$a_k(t, Y_t)' \equiv \mathbf{E}_t\left[\frac{J_{0,t,T}\hat{J}_{1,t,T}}{\mathbf{E}_t\left[J_{0,t,T}\hat{J}_{1,t,T}\phi(Z_t, Z_T)\right]}J_{2,k,t,T}\phi(Z_t, Z_T)\right]$$

$$b_k(t, Y_t)' \equiv \mathbf{E}_t\left[\frac{J_{0,t,T}\hat{J}_{1,t,T}}{\mathbf{E}_t\left[J_{0,t,T}\hat{J}_{1,t,T}\phi(Z_t, Z_T)\right]}\phi(Z_t, Z_T)[\hat{J}_{3,k,t,T} - \rho[\theta_k \circ G](Z_T)\mathcal{D}_{k,T}Z_T - \theta(Y_t)']\right]$$

where $J_{0,t,T}$, $J_{2,t,T}$ are defined above

$$\hat{J}_{1,t,v} \equiv \exp\{\rho \textstyle\int_t^v [\hat{h} \circ G](Z_s)ds\}$$

$$\hat{J}_{3,k,t,v} \equiv -\textstyle\int_t^v [(\sum_{j=1}^d \theta_j\partial\theta_j - \partial\hat{h}\sigma^Y) \circ G](Z_s)$$

and

$$\phi(Z_t, Z_v) \equiv \exp\left(-\rho(\Phi(Z_v) - \Phi(Z_t))\right).$$

By Ito's lemma the vector

$$\hat{X}'_{k,t,v} \equiv [J_{0,t,v}, \hat{J}_{1,t,v}, Z'_t, (\mathcal{D}_{k,t}Z_t)', J_{2,k,t,v}, \hat{J}_{3,k,t,v}]$$

satisfies the system of SDEs'

$$d\hat{X}_{k,t,v} = \hat{A}(\hat{X}_{k,t,v})dv + \sum_{j=1}^{d} \hat{B}_j dW_v^j \tag{3.13}$$

with $\hat{A} \in \mathcal{C}^1(\mathbb{R}^{2d+4})$ and $\hat{B}$ given by

$$\hat{A}(x) = \begin{bmatrix} -\rho x_1[h \circ G](x_3) \\ \rho x_2[\hat{h} \circ G](x_3) \\ \hat{\mu}(x_3) \\ \partial\hat{\mu}(x_3)x_4 \\ [\partial r\sigma^Y \circ G](x_3)x_4 \\ -[(\sum_{j=1}^{d}\theta_j\partial\theta_j - \partial\hat{h}\sigma^Y) \circ G](x_3)x_4 \end{bmatrix} \quad \text{and} \quad \hat{B}(x) = \begin{bmatrix} 0 \\ 0 \\ I_d \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

and initial condition $\hat{X}'_{k,t,t} = [1, 1, F(Y_t)', e'_k, 0, 0]$.

In this second representation the Wiener process enters additively into the system of SDEs. We can then approximate the martingale part of the diffusion equations without error by drawing multivariate Gaussian random variables. As we will show this increases the speed of convergence for the approximation of functions of the random variables $\hat{X}_{k,t,v}$.[2]

We have now expanded the space of state variables in such a way as to be able to write the functionals of Malliavin derivatives arising in the optimal portfolio policy as the terminal value of a system of SDEs. In principle any numerical solution scheme for SDEs could be applied to approximate these random variables. In what follows we will analyze the simplest scheme available, namely the

[2] If we would approximate the increments of the Wiener process by a convergent approximation like a binomial approximation the speed of convergence would not be improved. For the proof of increased speed of convergence it is therefore important to sample Gaussian random variables

Euler scheme. Our next two sections will establish that the Doss transformation increases the speed of convergence from $1/\sqrt{N}$ to $1/N$.

3.3 Approximating Functionals of Brownian Motion

Well-known results have established that that the speed of convergence of the Euler approximation of the solution of an SDE is of order $1/\sqrt{N}$. In this section we (i) extend the result to functional solutions of SDE and (ii) establish corresponding results for the Euler scheme combined with a Doss transformation of the underlying state variables. As we will show the transformation increases the speed of convergence to order $1/N$. We also provide asymptotic laws for the approximation errors.

3.3.1 Euler approximation without transformation

Consider a functional $\tilde{X}_{k,t,T}$ satisfying the SDE (3.12). In this section we demonstrate properties of its Euler approximation

$$\begin{aligned} \tilde{X}^N_{k,t,T} &= \tilde{X}_{k,t,t} + \sum_{n=0}^{N-1} \tilde{A}(\tilde{X}^N_{k,t,t+\frac{n(T-t)}{N}}) \left(\frac{T-t}{N}\right) \\ &+ \sum_{n=0}^{N-1}\sum_{j=1}^{d} \tilde{B}_j(\tilde{X}^N_{k,t,t+\frac{n(T-t)}{N}}) \left(W^j_{s+\frac{(n+1)(T-t)}{N}} - W^j_{s+\frac{n(T-t)}{N}}\right) \end{aligned} \tag{3.14}$$

From Kurtz and Protter (1991) and Jacod and Protter (1998) we can deduce the following result for the error distribution of the Euler approximation of the $2d+4\times d$ matrix $\tilde{X}_{t,T} = [\tilde{X}_{1,t,T}, \ldots, \tilde{X}_{d,t,T}]$

Theorem 36 *The asymptotic law of the estimator of the random matrix $\tilde{X}_{t,T}$ is*

$$\sqrt{N}(\tilde{X}^N_{t,T} - \tilde{X}_{t,T}) \Rightarrow \tilde{U}^{\tilde{X}}_{t,T}$$

where the error for the k^{th} column of $\tilde{X}_{t,T}$ is

$$U^{\tilde{X}}_{k,t,T} = -\frac{1}{\sqrt{2}}\tilde{\Omega}_{k,t,T}\int_t^T \tilde{\Omega}^{-1}_{k,t,v}\sum_{h,j=1}^{d}[\partial\tilde{B}_j\tilde{B}_h](\tilde{X}_{k,t,v})dB^{h,j}_v$$

where $[B^{h,j}]_{h,j\in\{k,\dots,d\}}$ is a $d^2 \times 1$ standard Brownian motion independent of W and

$$\tilde{\Omega}_{k,t,v} = \exp\left(\int_t^v [\partial\tilde{A} - \frac{1}{2}\sum_{j=1}^{d}\partial\tilde{B}_j(\tilde{B}_j)'](\tilde{X}_{k,t,s}ds + \sum_{j=1}^{d}\int_t^v \partial\tilde{B}_j(\tilde{X}_{k,t,s})dW^j_s\right)$$

Special cases of Theorem 36 provide the asymptotic laws of the errors involved in the estimation of the state variables Y, of their Malliavin derivatives $\mathcal{D}Y$, and of the functionals $H_i, i = 1,2,3$, and $J_i, i = 0,...,3$, appearing in the hedging components a and b. To state these results define the random variables

$$\tilde{\Omega}_{1,t,v} = \exp\left(-\rho\int_t^v h(Y_s)ds\right) \tag{3.15}$$

$$\tilde{\Omega}_{2,t,v} = \exp\left(-\frac{1}{2}\int_t^v \|\rho\theta(Y_s)\|^2 ds - \rho\sum_{j=1}^{d}\int_t^v \theta_j(Y_s)dW^j_s\right) \tag{3.16}$$

$$\tilde{\Omega}_{3,t,v} = \exp\left(\int_t^v [\partial\mu^Y - \frac{1}{2}\sum_{j=1}^{d}\partial\sigma^Y_{\cdot j}(\partial\sigma^Y_{\cdot j})'](Y_s)ds + \sum_{j=1}^{d}\int_t^v \partial\sigma^Y_{\cdot j}(Y_s)dW^j_s\right). \tag{3.17}$$

For the estimate of the state variable Y we obtain,

Corollary 37 *The asymptotic law of the estimator of the state variables Y is*

$$U^{\tilde{Y}^N}_t \equiv \sqrt{N}(\tilde{Y}^N_t - Y_t) \Rightarrow \tilde{U}^Y_t$$

where

$$\tilde{U}^Y_{t,v} = -\frac{1}{\sqrt{2}}\tilde{\Omega}_{3,t,v}\int_t^v \tilde{\Omega}^{-1}_{3,t,s}\sum_{h,j=1}^{d}[(\partial\sigma^Y_{\cdot j})\sigma^Y_{\cdot h}](Y_s)dB^{h,j}_s. \tag{3.18}$$

The law of the estimate of the Malliavin derivative of the state variable is,

Corollary 38 *The asymptotic law of the estimator of the Malliavin derivatives of the state variables $\mathcal{D}Y$ is*

$$U_{t,\cdot}^{\tilde{\mathcal{D}}_t^N Y} \equiv \sqrt{N}(\tilde{\mathcal{D}}_t^N Y_\cdot - \mathcal{D}_t Y_\cdot) \Rightarrow \tilde{U}_{t,\cdot}^{\mathcal{D}_t Y}$$

where the k^{th} column

$$\begin{aligned} \tilde{U}_{t,v}^{\mathcal{D}_{k,t}Y} &= \tilde{\Omega}_{3,t,v} \int_t^v \tilde{\Omega}_{3,t,s}^{-1} \{ (I_d \otimes (\mathcal{D}_{k,t} Y_s)') [\partial^2 \mu^Y (Y_s) ds + \sum_{j=1}^d \partial \sigma_{\cdot j}^Y (Y_s) dW_s^j] \tilde{U}_{t,s}^Y \quad (3.19) \\ & - \sum_{j=1}^d \partial \sigma_{\cdot j}^Y (Y_s)(I_d \otimes (\mathcal{D}_{k,t} Y_s)') \partial^2 \sigma_{\cdot j}(Y_s) \tilde{U}_{t,s}^Y ds \\ & - \frac{1}{\sqrt{2}} \sum_{h,j=1}^d [(I_d \otimes (\mathcal{D}_{k,t} Y_v)')[(\partial^2 \sigma_{\cdot j}^Y) \sigma_{\cdot h}^Y](Y_s) + [(\partial \sigma_{\cdot j}^Y) \partial \sigma_{\cdot h}^Y](Y_s) \mathcal{D}_{k,t} Y_s] dB_s^{h,j} \}. \end{aligned}$$

Our last corollary describes properties of estimators of the functionals H_i, J_i.

Corollary 39 *The asymptotic law of the estimator of the functional H_i, J_i are*

$$U_{t,T}^{\tilde{H}_i^N} := \sqrt{N}(\tilde{H}_{i,t,T}^N - H_{i,t,T}) \Rightarrow \tilde{U}_{t,T}^{H_i} \qquad \text{for } i = 1,2,3$$

$$U_{t,T}^{\tilde{J}_i^N} = \sqrt{N}(\tilde{J}_{i,t,T}^N - J_{i,t,T}) \Rightarrow \tilde{U}_{t,T}^{J_i} \qquad \text{for } i = 0,1$$

$$U_{t,T}^{\tilde{J}_{i,k}^N} = \sqrt{N}(\tilde{J}_{i,k,t,T}^N - J_{i,k,t,T}) \Rightarrow \tilde{U}_{t,T}^{J_{i,k}} \qquad \text{for } i = 2,3$$

where

$$\tilde{U}_{t,T}^{H_1} = J_{1,t,v} \tilde{U}_{t,T}^{J_0} + J_{0,t,v} \tilde{U}_{t,T}^{J_1}$$

$$\tilde{U}_{t,T}^{H_{2,k}} = J_{1,t,T} J_{2,k,t,T} \tilde{U}_{t,T}^{J_0} + J_{0,t,T} J_{2,k,t,T} \tilde{U}_{t,T}^{J_1} + J_{0,t,T} J_{1,t,T} \tilde{U}_{t,T}^{J_2}$$

$$\tilde{U}_{t,T}^{\tilde{H}_{3,k}} = J_{1,t,T}J_{3,k,t,T}\tilde{U}_{t,T}^{J_0} + J_{0,t,T}J_{3,k,t,T}\tilde{U}_{t,T}^{J_1} + J_{0,t,T}J_{1,t,T}\tilde{U}_{t,T}^{J_3}$$

($\tilde{U}_{t,T}^{H_{i,k}}$ denotes the k^{th} element of the vector $\tilde{U}_{t,T}^{H_i}, i = 1, 2$) and

$$\tilde{U}_{t,T}^{J_0} = -\tilde{\Omega}_{1,t,T}\int_t^T \tilde{\Omega}_{1,t,s}^{-1} J_{0,t,s}\partial h(Y_s)\tilde{U}_{t,s}^{Y} ds$$

$$\begin{aligned}\tilde{U}_{t,T}^{J_1} &= -\rho\tilde{\Omega}_{2,t,T}\int_t^T \tilde{\Omega}_{2,t,s}^{-1}\rho J_{1,t,s}\{[2\rho[\sum_{j=1}^{d}\theta_j\partial\theta_j](Y_s)ds \\ &\quad + \sum_{j=1}^{d}\partial\theta_j(Y_s)dW_s^j]\tilde{U}_{t,s}^{Y} \\ &\quad + \frac{1}{\sqrt{2}}\sum_{h,j=1}^{d}[\theta_j\theta_h - (\partial\theta_j)\sigma_{\cdot h}^{Y}](Y_s)dB_s^{h,j}\}\end{aligned}$$

$$\tilde{U}_{t,T}^{J_{2,k}} = \int_t^T [(\mathcal{D}_{k,t}Y_s)'\partial^2 r(Y_s)\tilde{U}_{t,s}^{Y} + \partial r(Y_s)\tilde{U}_{t,s}^{\mathcal{D}_{k,t}Y}]ds$$

$$\begin{aligned}\tilde{U}_{t,T}^{J_{3,k}} &= \int_t^T [(\mathcal{D}_{k,t}Y_s)'\partial\theta(Y_s)\tilde{U}_{t,s}^{Y} + \theta(Y_s)\tilde{U}_{t,s}^{\mathcal{D}_{k,t}Y}]ds \\ &\quad + \int_t^T \sum_{j=1}^{d}[(\mathcal{D}_{k,t}Y_s)'\partial^2\theta_j(Y_s)\tilde{U}_{t,s}^{Y} + \partial\theta_j(Y_s)\tilde{U}_{t,s}^{\mathcal{D}_{k,t}Y}]dW_s^j \\ &\quad - \frac{1}{\sqrt{2}}\sum_{h,j=1}^{d}[(\mathcal{D}_{k,t}Y_s)'[(\partial^2\theta_j)\sigma_{\cdot h}^{Y}] + [\partial\theta_j\partial\sigma_{\cdot h}^{Y}](Y_s)]dB_s^{h,j}.\end{aligned}$$

In all these expressions $\tilde{U}^{Y}$ and $\tilde{U}^{\mathcal{D}_{k,t}Y}$ are given in (3.18) and (3.19), respectively.

3.3.2 Euler approximation with Doss transformation

We now establish the corresponding results with Doss transformation. We consider the Euler approximation obtained after application of a Doss transformation to the state variables.

$$\hat{X}^N_{k,t,T} = \hat{X}_{k,t,t} + \sum_{n=0}^{N-1} \hat{A}(\hat{X}^N_{k,t,t+\frac{n(T-t)}{N}}) \left(\frac{T-t}{N}\right) + \sum_{n=0}^{N-1}\sum_{j=1}^{d} \hat{B}_j \left(W^j_{s+\frac{(n+1)(T-t)}{N}} - W^j_{s+\frac{n(T-t)}{N}}\right) \tag{3.20}$$

The error distribution of the Euler approximation of the $d \times d$ matrix $\hat{X}_{t,T} = [\hat{X}_{1,t,T}, \ldots, \hat{X}_{d,t,T}]$ is given first.

Theorem 40 *The asymptotic law of the estimator of the random matrix $\hat{X}_{t,T}$ is*

$$N(\hat{X}^N_{t,T} - \hat{X}_{t,T}) \Rightarrow \hat{U}^{\hat{X}}_{t,T}$$

where

$$U^{\hat{X}}_{k,t,T} = -\hat{\Omega}_{k,t,T} \int_t^T \hat{\Omega}^{-1}_{k,t,v} \partial \hat{A}(\hat{X}_{k,t,v})[\frac{1}{2} dX_{k,t,v} + \frac{1}{\sqrt{12}} \sum_{j=1}^{d} \hat{B}_j dB^j_v]$$

where $[B^j]_{j\in\{1,\ldots,d\}}$ is a $d \times 1$ standard Brownian motion independent of W and $B^{h,j}$ and

$$\hat{\Omega}_{k,t,v} = \exp\left(\int_t^v \partial \hat{A}(\hat{X}_{k,t,s}) ds\right)$$

This theorem shows that the speed of convergence increases if we use the Doss transformation. It also highlights the fact that the limit law is different and involves exponentials of a bounded total variation process instead of a stochastic integral.

In order to specialize the results of this theorem we introduce the random variables:

$$\hat{\Omega}_{1,t,v} = \exp\left(-\rho \int_t^v h(Y_s) ds\right)$$

$$\hat{\Omega}_{2,t,v} = \exp\left(\rho \int_t^v \hat{h}(Y_s) ds\right)$$

$$\hat{\Omega}_{3,t,v} = \exp\left(\int_t^v [(\partial\hat{m})\sigma^Y](Y_s)ds\right).$$

With these definitions the law of the error of the estimate of the state variable is given by the following corollary.

Corollary 41 *The asymptotic law of the estimator of the state variable Y is given by*

$$U_{t,T}^{\hat{Y}^N} \equiv N(\hat{Y}_T^N - Y_T) \Rightarrow \hat{U}_{t,T}^Y$$

where

$$\hat{U}_{t,T}^Y = -\sigma^Y(Y_T)\hat{\Omega}_{3,t,T}\int_t^T \hat{\Omega}_{3,t,s}^{-1}\partial\hat{m}(Y_s)\left[\frac{1}{2}(\hat{m}(Y_s)ds + dW_s) + \frac{1}{\sqrt{12}}dB_s\right]. \quad (3.21)$$

For the Malliavin derivative of the state variables we get,

Corollary 42 *The asymptotic law of the estimator of the Malliavin derivatives of the state variables without the transformation is*

$$U_{t,\cdot}^{\hat{\mathcal{D}}_t^N Y} \equiv N(\hat{\mathcal{D}}_t^N Y_\cdot - \mathcal{D}_t Y_\cdot) \Rightarrow \hat{U}_{t,\cdot}^{\mathcal{D}_t Y}$$

where

$$\hat{U}_{t,T}^{\mathcal{D}_t Y} = \sigma^Y(Y_T)\hat{U}_{t,T}^{\mathcal{D}} \otimes I_d)\hat{U}_{t,T}^Y \quad (3.22)$$

and where the k^{th} column of $\hat{U}_{t,T}^{\mathcal{D}_t Z}$ is

$$\begin{aligned}\hat{U}_{t,T}^{\mathcal{D}_{k,t} Z} &= \hat{\Omega}_{3,t,T}\int_t^T (\hat{\Omega}_{3,t,s})^{-1}[(I_d \otimes (\sigma^Y(Y_s)^{-1}\mathcal{D}_{k,t}Y_s)') \\ &\quad \times[(\partial\hat{m} \otimes I_d)(\partial\sigma^Y)' + (\sigma^Y)'\partial^2\hat{m}](Y_s)[\hat{U}_{t,s}^Y ds - \frac{1}{2}[\sigma^Y(\hat{m}(Y_s)ds + dW_s)]] \\ &\quad +[\partial r\sigma^Y](Y_s)[\hat{U}^{\mathcal{D}_{k,t}Z}ds - \frac{1}{2}\partial\hat{m}(Y_s)\mathcal{D}_{k,t}Y_s ds]]\end{aligned}$$

Similarly, for the asymptotic law of the approximated random variables $\hat{H}_{k,t,T}$ involved in the calculation of the functions a and b we obtain:

Corollary 43 *The asymptotic law of the estimator of the functional $\widehat{H}_i$, $\widehat{J}_i$ are*

$$U_{t,T}^{\hat{H}_i^N} \equiv N(\hat{H}_{i,t,T}^N - H_{i,t,T}) \Rightarrow \hat{U}_{t,T}^{H_i} \qquad \text{for } i = 1,2,3$$

$$U_{t,T}^{\hat{J}_i^N} \equiv N(\hat{J}_{i,t,T}^N - J_{i,t,T}) \Rightarrow \hat{U}_{t,T}^{J_i} \qquad \text{for } i = 0,1$$

$$U_{t,T}^{\hat{J}_{i,k}^N} \equiv N(\hat{J}_{i,k,t,T}^N - J_{i,k,t,T}) \Rightarrow \hat{U}_{t,T}^{J_{i,k}} \qquad \text{for } i = 2,3$$

where

$$\hat{U}_{t,T}^{H_1} = J_{1,t,v} U_{t,T}^{J_0} + J_{0,t,v} U_{t,T}^{J_1}$$

$$\hat{U}_{t,T}^{H_{2,k}} = J_{1,t,T} J_{2,k,t,T} \hat{U}_{t,T}^{J_0} + J_{0,t,T} J_{2,t,T} \hat{U}_{t,T}^{J_1} + J_{0,t,T} J_{1,t,T} \hat{U}_{t,T}^{J_2}$$

$$\hat{U}_{t,T}^{H_{3,k}} = J_{1,t,T} J_{3,k,t,T} \hat{U}_{t,T}^{J_0} + J_{0,t,T} J_{3,t,T} \hat{U}_{t,T}^{J_1} + J_{0,t,T} J_{1,t,T} \hat{U}_{t,T}^{J_3}.$$

($\hat{U}_{t,T}^{H_{i,k}}$ denotes the k^{th} element of the vector $\tilde{U}^{H_i}, i = 2,3$) and

$$\begin{aligned} U_{t,T}^{\hat{J}_0} &= -\widehat{\Omega}_{1,t,T} \int_t^T \widehat{\Omega}_{1,t,s}^{-1} [\rho J_{0,t,s} (\frac{1}{2}([(\partial h) \sigma^Y \hat{m}](Y_s) ds + [(\partial h) \sigma^Y](Y_s) dW_s) \\ &\quad -\partial h(Y_s) \hat{U}_{t,s}^Y ds) - \frac{1}{2} dJ_{0,t,s}] \end{aligned}$$

$$\hat{U}_{t,T}^{J_1} = \frac{J_{1,t,T}}{\hat{J}_{1,t,T}} \hat{U}_{t,T}^{\hat{J}_1} - \rho J_{1,t,T} \theta(Y_T)' \hat{U}_{t,T}^Y$$

$$\begin{aligned} \hat{U}_{t,T}^{J_{2,k}} &= \int_t^T (\sigma^Y(Y_s)^{-1} \mathcal{D}_{k,t} Y_s)' [[(\partial r \otimes I_d)(\partial \sigma^Y)' + (\sigma^Y)' \partial^2 r](Y_s) [\hat{U}_{t,s}^Y \\ &\quad -\frac{1}{2}(\hat{m}(Y_s) ds + dW_s)]] + [\partial r \sigma^Y](Y_s) [\hat{U}_{t,s}^{\mathcal{D}_{k,t} Z} - \frac{1}{2} d\mathcal{D}_{k,t} Z_s] \end{aligned}$$

$$\begin{aligned} \hat{U}_{t,T}^{J_{3,k}} &= U_{t,T}^{\hat{J}_3} - \rho[\theta'(\sigma^Y)^{-1}](Y_T) \hat{U}_{t,T}^{\mathcal{D}_{k,t} Y} \\ &\quad -\rho(\hat{U}_{t,T}^Y)' [((\sigma^Y)')^{-1} \partial \theta + (\theta'(\sigma^Y)^{-1} \otimes ((\sigma^Y)')^{-1}) \partial (\sigma^Y)'] \mathcal{D}_{k,t} Y_T \end{aligned}$$

with

$$\begin{aligned} U_{t,T}^{\hat{J}_1} &= \hat{\Omega}_{2,t,T} \int_t^T \hat{\Omega}_{2,t,s}^{-1} [\rho \hat{J}_{1,t,s} (\frac{1}{2}([(\partial \hat{h})\sigma^Y \hat{m}](Y_s) ds + [(\partial \hat{h})\sigma^Y](Y_s) dW_s) \\ &\quad - \partial \hat{h}(Y_s) \hat{U}_{t,s}^Y ds) - \frac{1}{2} d\hat{J}_{1,t,s}] \end{aligned}$$

$$\begin{aligned} \hat{U}_{t,T}^{\hat{J}_{3,k}} &= \int_t^T [(\sigma^Y(Y_s)^{-1} \mathcal{D}_{k,t} Y_s)' [\sum_{j=1}^d ((\partial\theta_j)'\partial\theta_j + \theta_j \partial^2 \theta_j)(\sigma^Y)^{-1} \\ &\quad - ((\partial \hat{h} \otimes I_d)(\partial \sigma^Y)' - (\sigma^Y)' \partial^2 \hat{h})](Y_s)[\hat{U}_{t,s}^Y ds - \frac{1}{2}\sigma^Y(Y_s)(\hat{m}(Y_s) ds + dW_s)] \\ &\quad - [\sum_{j=1}^d \theta_j \partial\theta_j - \partial \hat{h} \sigma^Y](Y_s)[\hat{U}_{t,s}^{\mathcal{D}_{k,t} Z} ds + \frac{1}{2} d\mathcal{D}_{k,t} Z_s]] \end{aligned}$$

and where $\hat{U}^Y$ and $\hat{U}^{\mathcal{D}_{k,t}Z}$ are defined in (3.21) and (3.22), respectively.

3.4 Expected Approximation Errors

In this section we analyze the expected approximation errors. We show that the order of convergence of expected approximation errors is identical for both methods and explain the source of this result. Our analysis is based on weak convergence results for the drift and martingale part of the SDE. We derive expected approximation errors for both schemes in the form of expectations of known random variables. Our findings can be viewed as probabilistic counterparts of the results of Talay and Tubaro (1991) and Bally and Talay (1996) who have characterized the expected approximation error as the solution of a PDE.

3.4.1 Expected Approximation Error: General Results

Throughout this section we assume that $X_{t,v}, v \in [t,T]$ satisfies (3.12). Our first result describes the convergence of the expected approximation error using the Euler scheme.

Theorem 44 *For $g \in \mathcal{C}^1(\mathbb{R}^d)$ such that*

$$\lim_{r\to\infty} sup_N \mathbf{E}\left[\mathbf{1}_{\{\|\tilde{X}^N_{t,T}\|>r\}} g(\tilde{X}^N_{t,T})|\mathcal{F}_t\right] = 0 \tag{3.23}$$

P-*a.s. we have*

$$N\mathbf{E}\left[g(\tilde{X}^N_{t,T}) - g(X_{t,T})|\mathcal{F}_t\right] \to \frac{1}{2}\tilde{K}_{t,T}(X_t) \tag{3.24}$$

where

$$\begin{aligned}\tilde{K}_{t,T}(X_t) \quad = \quad -\mathbf{E}\Bigg[\partial g(X_{t,T})\tilde{\Omega}_{t,T}\int_t^T \tilde{\Omega}^{-1}_{t,s}\{\sum_{j=1}^d[\partial B_j A](X_{t,s})dW^j_s \qquad (3.25)\\ + [\partial A - \sum_{j=1}^d \partial\tilde{B}_j\tilde{B}_j](X_{t,s})dX_{t,s}\}|\mathcal{F}_t\Bigg]\end{aligned}$$

with $\tilde{\Omega}_{t,v} = \exp\left(\int_t^v[\partial A(X_{t,s}) - \frac{1}{2}\sum_{j=1}^d(\partial B_j)^2]ds + \sum_{j=1}^d\int_t^v \partial B_j dW^j_s\right)$.

This Theorem provides a probabilistic counterpart to the results of Talay and Tubaro (1991) and Bally and Talay (1996) who characterized the second order expected approximation error of the Euler scheme as the solution of a PDE. Our proof uses probabilistic arguments to provide a representation in the form of a conditional expectation of a random variable whose components depend on the coefficients of the SDE satisfied by $X_{t,v}$.

Note that the second order bias $\tilde{K}_{t,T}(X_t)$ can be approximated by simulation. Our result could then be used to develop approximation schemes that correct for second order bias. As will be discussed further such schemes are preferable as soon as the number of state variables is sufficiently large to make the numerical solution of PDEs costly in terms of computation.

The corresponding result for the scheme based on the transformed state variables is given by the following theorem.

Theorem 45 *Suppose that Assumption A holds. For $f \in C^1(\mathbb{R}^d)$ such that*

$$\lim_{r\to\infty} sup_N \mathbf{E}\left[\mathbf{1}_{\{\|\hat{X}^N_{t,T}\|>r\}} g(\hat{X}^N_t)|\mathcal{F}_t\right] = 0 \tag{3.26}$$

$\mathbf{P}$-a.s. we have

$$N\mathbf{E}\left[g(\hat{X}^N_{t,T}) - g(X_{t,T})|\mathcal{F}_t\right] \to \frac{1}{2}\hat{K}_{t,T}(X_t) \tag{3.27}$$

where

$$\begin{aligned}\hat{K}_{t,T}(X_t) &= -\mathbf{E}\left[[\partial g\sigma](X_{t,T})\hat{\Omega}_{t,T}\right. \\ &\left.\times \int_t^T \hat{\Omega}^{-1}_{t,s}\left([(\partial\hat{m})\sigma\hat{m}](X_{t,s})ds + [(\partial\hat{m}\sigma)](X_{t,s})dW_s\right)|\mathcal{F}_t\right]\end{aligned} \tag{3.28}$$

with $\hat{\Omega}_{t,T} = \exp\left(\int_t^v [(\partial\hat{m})\sigma](X_{t,s})ds\right)$.

A comparison of (3.25) with (3.28) suggests that it will be difficult, in general, to establish the dominance of one method over the other. Indeed, as the asymptotic expressions show the speed of convergence crucially depends of the volatility coefficient σ.

The difference in the rates of convergence for the limit distribution and the expected approximation error can be explained as follows. As we have seen the asymptotic law of the approximation error using the Euler scheme is the product of two random variables that are independent. Since the second of these is a stochastic integral with null expectation the asymptotic law is centered around zero. The expected approximation error for the Euler scheme can therefore not be of order $1/\sqrt{N}$. In contrast, the expectation of the random variable describing the asymptotic law of the Euler approximation with Doss transformation differs from zero. This approximation scheme has therefore a second order bias, which implies convergence speed of order $1/N$.

3.4.2 Expected Approximation Error without Transformation

We now specialize these results to the random variables arising in the hedging components of the optimal portfolio. The expected approximation errors identified will be used in the next section to find the asymptotic distribution of the Monte Carlo estimators of the functions a and b.

In order to state these results define the random variables $\tilde{V}_{t,T}^{Y}$ and $\tilde{V}_{t,T}^{\mathcal{D}_{k,t}Y}$,

$$2\tilde{V}_{t,T}^{Y} = -\tilde{\Omega}_{3,t,T}\int_t^T \tilde{\Omega}_{3,t,s}^{-1}\left[[\partial\mu^Y - \sum_{j=1}^{d}\partial\sigma_{\cdot j}^{Y}\partial\sigma_{\cdot j}^{Y}](Y_s)dY_s + \sum_{j=1}^{d}[\partial\sigma_{\cdot j}^{Y}\mu^Y](Y_s)dW_s^j\right]$$

$$\begin{aligned}
2\tilde{V}_{t,T}^{\mathcal{D}_{k,t}Y} &= \tilde{\Omega}_{3,t,T}\int_t^T \tilde{\Omega}_{t,s}^{-1}\{(I_d\otimes(\mathcal{D}_{k,t})')[\partial^2\mu^Y(Y_s)ds \\
&\quad + \sum_{j=1}^{d}\partial^2\sigma_{\cdot j}^{Y}dW_s^{hj}]2\tilde{V}_{t,s}^{Y}[(I_d\otimes(\mathcal{D}_{k,t}Y_s)')[\partial^2\mu^Y - \sum_{j=1}^{d}\partial^2\sigma_{\cdot j}^{Y}\partial\sigma_{\cdot j}^{Y}](Y_s) \\
&\quad - \sum_{j=1}^{d}\partial\sigma_{j}^{Y}(Y_s)(I_d\otimes(\mathcal{D}_{k,t}Y_s)')\partial^2\sigma_{\cdot j}^{Y}(Y_s)]dY_s \\
&\quad + [\partial\mu^Y - \sum_{j=1}^{d}\sigma_{\cdot j}^{Y}\sigma_{\cdot j}^{Y}](Y_s)d\mathcal{D}_{k,t}Y_s \\
&\quad + [(I_d\otimes(\mathcal{D}_{k,t}Y_s)')\sum_{j=1}^{d}[\partial^2\sigma_{\cdot j}^{j}\mu^Y](Y_s) + \sum_{j=1}^{d}[\partial\sigma_{\cdot j}^{Y}(Y_s)\mathcal{D}_{k,t}Y_s]dW_s^j \\
&\quad + (I_d\otimes(\mathcal{D}_{k,t}Y_s)')[\sum_{j=1}^{d}\partial^2\sigma_{\cdot j}^{Y}(\partial\mu^Y - \sum_{h=1}^{d}\partial\sigma_{\cdot h}^{Y}\partial\sigma_{\cdot h}^{Y})\sigma_{\cdot j}^{Y}](Y_s)ds \\
&\quad + (I_d\otimes(\mathcal{D}_{k,t}Y_s)')\sum_{j=1}^{d}[\partial^2\sigma_{\cdot j}^{Y}\partial\sigma_{\cdot j}^{Y}\mu^Y](Y_s)ds \\
&\quad - \sum_{j=1}^{d}\sigma_{\cdot j}^{Y}(Y_s)(I_d\otimes(\mathcal{D}_{k,t}Y_s)')\partial^2\sigma_{\cdot j}^{Y}(Y_s)ds
\end{aligned}$$

Our first corollary describes the expected approximation error for the functionals $H_{i,k,t,T}, i = 1, 2, 3$.

Corollary 46 *The asymptotic law of the conditional expectation of the functional* $H_{i,k,t,T}$ *is*

$$N\mathbf{E}\left[\tilde{H}^N_{1,t,T} - H_{1,t,T}|\mathcal{F}_t\right] \to \tfrac{1}{2}\tilde{K}^{H_1}_{t,T}(X_t)$$

$$N\mathbf{E}\left[\tilde{H}^N_{i,k,t,T} - H_{i,k,t,T}|\mathcal{F}_t\right] \to \tfrac{1}{2}\tilde{K}^{H_{i,k}}_{t,T}(X_t) \qquad \textit{for } i = 2,3$$

where

$$\tilde{K}^{H_1}_{t,T}(X_t) = \mathbf{E}[J_{1,t,v}V^{J_0}_{t,T} + J_{0,t,v}V^{J_1}_{t,T}|\mathcal{F}_t]$$

$$\tilde{K}^{H_{2,k}}_{t,T}(X_t) = \mathbf{E}[J_{1,t,T}J_{2,k,t,T}\tilde{V}^{J_0}_{t,T} + J_{0,t,T}J_{2,t,T}\tilde{V}^{J_1}_{t,T} + J_{0,t,T}J_{1,t,T}\tilde{V}^{J_{2,k}}_{t,T}|\mathcal{F}_t]$$

$$\tilde{K}^{H_{3,k}}_{t,T}(X_t) = \mathbf{E}[J_{1,t,T}J_{3,k,t,T}\tilde{V}^{J_0}_{t,T} + J_{0,t,T}J_{3,t,T}\tilde{V}^{J_1}_{t,T} + J_{0,t,T}J_{1,t,T}\tilde{V}^{J_{3,k}}_{t,T}|\mathcal{F}_t].$$

In these expressions

$$2\tilde{V}^{J_0}_{t,T} = -\Omega_{1,t,T}\int_t^T \Omega^{-1}_{1,t,s}[J_{0,t,s}\partial h(Y_s)2\tilde{V}^Y_{t,s}ds - h(Y_s)dJ_{0,t,s} + J_{0,t,s}(\partial h(Y_s))dY_s]$$

$$\begin{aligned}
2\tilde{V}^{J_{1,t}}_{t,T} \quad = \quad & -\tilde{\Omega}_{2,t,T}\int_t^T \tilde{\Omega}^{-1}_{2,t,s}\Bigg\{\rho J_{1,t,s}[\rho[\theta'\partial\theta](Y_s)ds + \sum_{j=1}^d \partial\theta_j(Y_s)dW^j_s]2\tilde{V}^Y_{t,s} \\
& +\frac{3}{2}\|\rho\theta(Y_s)\|^2 dJ_{1,t,s} + \rho J_{1,t,s}[2\rho\theta'\partial\theta - \sum_{j=1}^d \partial\theta_j\sigma^Y_{\cdot j}]dY_s \\
& -\frac{1}{2}\rho J_{1,t,s}\|\rho\theta(Y_s)\|^2\sum_{j=1}^d[\theta_j - \partial\theta_j\mu^Y](Y_s)dW^j_s \\
& +\frac{1}{2}\rho\sum_{j=1}^d[\partial\theta_j\sigma^Y_j\mu^Y](Y_s) - \frac{1}{2}\rho^2[2\rho\sum_{j=1}^d[\theta_j\partial\theta_j\mu^Y](Y_s)dJ_{1,t,s} \\
& +\rho J_{1,t,s}\sum_{j=1}^d[\theta_j(\partial\mu^Y - \sum_{h=1}^d \partial\sigma^Y_{\cdot h}\partial\sigma^Y_{\cdot h})\sigma^Y_{\cdot j}](Y_s)ds \\
& +\frac{1}{2}\rho^2 J_{1,t,s}\|\rho\theta(Y_s)\|^2\sum_{j=1}^d[\partial\theta_j\mu^Y](Y_s)ds \\
& -\rho^2 J_{1,t,s}\sum_{j=1}^d[\theta_j\partial\theta_j](2\tilde{V}^Y_{t,s})ds\Bigg\}
\end{aligned}$$

$$2\tilde{V}_{t,T}^{J_{2,k}} = \rho \int_t^T [[(\mathcal{D}_{k,t}Y_s)'\partial^2 r(Y_s)[2\tilde{V}_{t,s}^{Y} + dY_s] + \partial r(Y_s)[2\tilde{V}_{t,s}^{\mathcal{D}_{k,t}Y} ds + d\mathcal{D}_{k,t}Y_s]]$$

and

$$\begin{aligned} 2\tilde{V}_{t,T}^{J_{3,k}} &= \int_t^T (\mathcal{D}_{k,t}Y_s)'[\sum_{j=1}^{d} [\partial^2\theta_j \partial\sigma_{\cdot j}^{Y} \mu^{Y}](Y_s)ds \\ &+ \sum_{j=1}^{d} \partial^2\theta_j (\partial\mu^{Y} - \sum_{h=1}^{d} [\partial\sigma_{\cdot h}^{Y} \partial\sigma_{\cdot h}^{Y})\sigma_{\cot j}^{Y}](Y_s)ds \\ &-[\partial^2 h - \rho\partial^2 r](Y_s) + \sum_{j=1}^{d} [\partial^2\theta_j \sigma_{\cdot j}^{Y}](Y_s)dY_s \\ &-([\partial^2 h - \rho\partial^2 r](Y_s)ds - \sum_{j=1}^{d} \partial^2\theta_j (Y_s)dW_s^j)2\tilde{V}_{t,s}^{Y} \\ &+ \sum_{j=1}^{d} \partial\theta_j (Y_s)[[\partial\mu^{Y} - \sum_{h=1}^{d} \partial\sigma_{\cdot h}^{Y} \partial\sigma_{\cdot h}^{Y}](Y_s)\partial\sigma \cdot j^{Y}(Y_s)\mathcal{D}_{k,t}Y_s ds \\ &+(I_d \otimes (\mathcal{D}_{k,t}Y_s)')[[\partial^2\sigma_{\cdot j}^{Y}\mu^{Y}](Y_s) + [\partial^2\mu^{Y} - \sum_{h=1}^{d} \partial\sigma_{\cdot h}^{Y}\partial\sigma_{\cdot h}^{Y}]](Y_s)ds \\ &+\partial^2\sigma_{\cdot j}^{Y}(Y_s)dY_s - \sum_{h=1}^{d} \sigma_{\cdot h}^{Y}(I_d \otimes (\mathcal{D}_{k,t}Y_s)')\sigma_{\cdot j}^{Y} ds + [\sigma_{\cdot j}^{Y}\mu^{Y}](Y_s)ds] \\ &-[\partial h - \rho\partial r + \sum_{j=1}^{d} \partial\theta_j \partial\sigma_{\cdot j}^{Y}](Y_s)d\mathcal{D}_{k,t}Y_s \\ &-([\partial h - \rho\partial r](Y_s)ds + \sum_{j=1}^{d} \partial\theta_j (Y_s)dW_s^j)2\tilde{V}_{t,s}^{\mathcal{D}_{k,t}Y} \end{aligned}$$

3.4.3 Expected Approximation Error with Transformation.

We now present the expected approximation error for the estimates of $\hat{H}_{1,t,T}^{N}$, $\hat{H}_{2,k,t,T}^{N}$ and $\hat{H}_{3,k,t,T}^{N}$. The results can be derived immediately from corollary (43) using the fact that stochastic integrals with respect to the independent Brownian motion processes B are zero.

Define the random variables

$$\hat{V}^{\mathcal{D}_t Y} = \sigma^Y(Y_T)\hat{V}^{\mathcal{D}_t Z} + (\mathcal{D}_{k,t}Y_T)'((\sigma^Y(Y_T))' \otimes I_d)\hat{V}^Y_T$$

$$\begin{aligned} \hat{V}^{\mathcal{D}_t Z}_{t,T} \;=\;& \hat{\Omega}_{3,t,v}\int_t^v (\hat{\Omega}_{3,t,s})^{-1}[((\mathcal{D}_{k,t}Y_T)'(\sigma^Y(Y_T))' \otimes I_d)[(I_d \otimes (\sigma^Y)')\partial^2 \hat{m} \\ & +(\partial\hat{m} \otimes I_d)\partial(\sigma^Y)'](Y_s)((2\hat{V}^Y_{t,s} - [\sigma^Y \hat{m}](Y_s))ds - \sigma^Y(Y_s)dW_s) \\ & -[(\partial\hat{m})\sigma^Y(\partial\hat{m})](Y_s)\mathcal{D}_{k,t}Y_s ds] \end{aligned}$$

$$\hat{V}^Y_{t,T} = -\sigma^Y(Y_T)\hat{\Omega}_{3,t,T}\int_t^T \hat{\Omega}^{-1}_{3,t,T}\partial\hat{m}(Y_s)[\hat{m}(Y_s)ds + dW_s].$$

We have,

Corollary 47 *The asymptotic law of the conditional expectation of the functional* $H_{i,k,t,T}$ *is*

$$N\mathbf{E}[\hat{H}^N_{1,t,T} - H_{1,t,T}|\mathcal{F}_t] \to \tfrac{1}{2}\hat{K}^{H_1}_{t,T}(X_t)$$

$$N\mathbf{E}[\hat{H}^N_{i,k,t,T} - H_{i,k,t,T}|\mathcal{F}_t] \to \tfrac{1}{2}\hat{K}^{H_{i,k}}_{t,T}(X_t) \qquad \text{for } i = 2, 3$$

where

$$\hat{K}^{H_1}_{t,T}(X_t) = \mathbf{E}[J_{1,t,v}V^{J_0}_{t,T} + J_{0,t,v}V^{J_1}_{t,T}|\mathcal{F}_t]$$

$$\hat{K}^{H_{2,k}}_{t,T}(X_t) = \mathbf{E}[J_{1,t,T}J_{2,k,t,T}\hat{V}^{J_0}_{t,T} + J_{0,t,T}J_{2,t,T}\hat{V}^{J_1}_{t,T} + J_{0,t,T}J_{1,t,T}\hat{V}^{J_2}_{t,T}|\mathcal{F}_t]$$

$$\hat{K}^{H_{3,k}}_{t,T}(X_t) = \mathbf{E}[J_{1,t,T}J_{3,k,t,T}\hat{V}^{J_0}_{t,T} + J_{0,t,T}J_{3,t,T}\hat{V}^{J_1}_{t,T} + J_{0,t,T}J_{1,t,T}\hat{V}^{J_3}_{t,T}|\mathcal{F}_t].$$

In these expressions

$$V^{\hat{J}_0}_{t,T} = \Omega_{1,t,T}\int_t^T \Omega^{-1}_{1,t,s}[J_{0,t,s}(([(\partial h)\sigma^Y\hat{M}](Y_s)ds + [(\partial h)\sigma^Y](Y_s)dW_s) - \partial h(Y_s)\hat{V}^Y_{t,s}ds) - dJ_{0,t,s}]$$

$$\hat{V}^{J_1}_{t,T} = \frac{J_{1,t,T}}{\hat{J}_{1,t,T}}\hat{V}^{\hat{J}_1}_{t,T} - \rho J_{1,t,T}\theta(Y_T)'\hat{V}^Y_{t,T}$$

with

$$V_{t,T}^{\hat{J}_1} = \hat{\Omega}_{2,t,T} \int_t^T \hat{\Omega}_{2,t,s}^{-1} [\hat{J}_{1,t,s}(([(\partial\hat{h})\sigma^Y \hat{M}](Y_s)ds + [(\partial\hat{h})\sigma^Y](Y_s)dW_s) - \partial\hat{h}(Y_s)\hat{V}_{t,s}^Y ds) - d\hat{J}_{1,t,s}]$$

$$\begin{aligned}\hat{V}_{t,T}^{J_{2,k}} &= \rho \int_t^T [(\mathcal{D}_{k,t}Y_s)'[\partial^2 r + ((\sigma^Y)')^{-1}(\partial r \otimes I_d)\partial(\sigma^Y)'](Y_s)(2\hat{V}_{t,s}^Y ds \\ &\quad -[\sigma^Y \hat{m}](Y_s)ds - \sigma^Y(Y_s)dW_s) \\ &\quad -\partial r(Y_s)(2\hat{V}_{t,s}^{\mathcal{D}_{k,t}Y} ds - [\sigma^Y \partial\hat{m}](Y_s)\mathcal{D}_{k,t}Y_s ds)]\end{aligned}$$

$$\begin{aligned}\hat{V}_{t,T}^{J_{3,k}} &= V_{t,T}^{\hat{J}_3} - \rho[\theta'(\sigma^Y)^{-1}](Y_T)\hat{V}_{t,T}^{\mathcal{D}_{k,t}Y} \\ &\quad -\rho(\hat{V}_{t,T}^Y)'[((\sigma^Y)')^{-1}\partial\theta + (\theta'(\sigma^Y)^{-1} \otimes ((\sigma^Y)')^{-1})\partial(\sigma^Y)']\mathcal{D}_{k,t}Y_T\end{aligned}$$

$$\begin{aligned}d\hat{V}_{t,T}^{\hat{J}_{3,k}} &= \int_t^T [(\mathcal{D}_{k,t}Y_s)'[\partial^2 h - \rho\partial^2 r](Y_s) \\ &\quad +((\partial h + \partial\hat{h} - \rho\partial r) \otimes I_d)\partial(\sigma^Y)'](Y_s)(2\hat{V}_{t,s}^Y ds \\ &\quad -[\sigma^Y \hat{m}](Y_s) - \sigma^Y(Y_s)dW_s) \\ &\quad -[\partial h + \partial\hat{h} - \rho\partial r](Y_s)(2V_{t,s}^{\mathcal{D}_{t,k}Y} ds - [\sigma^Y \partial\hat{m}](Y_s)\mathcal{D}_{k,t}Y_s ds)].\end{aligned}$$

We now have all the necessary results to establish the asymptotic distribution of the approximation error for hedging terms using our Monte Carlo method.

3.5 Asymptotic Laws of Hedging Terms Estimators

We now use results from the previous sections to derive the asymptotic distribution of the approximation errors of the hedging terms. As we have seen this problem can be embedded in the more general problem of finding the conditional expectation of a function of the terminal value of an SDE. When the finite dimensional distribution of the solution of an SDE is unknown we can obtain an estimator of the expected value by sampling independent replications of the numerical solutions of the SDE. The overall approximation error of this numerical scheme should then be analyzed. From Duffie and Glynn (1995) it is known that this error has two components. The first is the expected approximation error we derived in the previous section. The second is related to the approximation error of the conditional expectation relying on a law of large numbers for independent random variates. Since we have found an explicit expression for the expected approximation error we can completely characterize the asymptotic distribution of the overall procedure.

In this section we will first derive a general result that will be applied to our context. We then derive an estimator of functions of terminal values of an SDE that are asymptotically equivalent to infeasible estimators we would have been obtained if we could sample from the true distribution of these random variates.

3.6 Asymptotic Laws of Hedging Terms Estimators

We now use results from the previous sections to derive the asymptotic distribution of the approximation errors of the hedging terms. As we have seen this problem can be embedded in the more general problem of finding the conditional expectation of a function of the terminal value of an SDE. When the finite dimensional distribution of the solution of an SDE is unknown we can obtain an estimator of the expected value by sampling independent replications of the numerical solutions of the SDE. The overall approximation error of this numerical scheme should then be analyzed. From Duffie and Glynn (1995) it is known that this error has two components. The first is the expected approximation error we derived in the previous section. The second is related to the approximation error of the conditional expectation relying on a law of large numbers for independent random variates. Since we have found an explicit expression for the expected approximation error we can completely characterize the asymptotic distribution of the overall procedure.

In this section we will first derive a general result that will be applied to our context. We then derive an estimator of functions of terminal values of an SDE that are asymptotically equivalent to infeasible estimators we would have been obtained if we could sample from the true distribution of these random variates.

3.6.1 General Results: Asymptotic Laws of Monte Carlo Estimators.

Suppose that we wish to calculate $\mathbf{E}[g(X_{t,T})|\mathcal{F}_t]$ where $X_{t,T}$ solves

$$dX_{t,v} = A(X_{t,v})dv + \sum_{j=1}^{d} B_j(X_{t,v})dW_v^j.$$

Our next result describes the asymptotic laws for the scheme with and without transformation,

Theorem 48 *Under the conditions of theorem 44 and for* $g \in \mathcal{C}^1(\mathbb{R}^d)$ *such that* $g(X_{t,T}) \in \mathbb{D}^{1,2}$ *we have*

$$\sqrt{M}\left(\frac{1}{M}\sum_{i=1}^{M} g(\tilde{X}_{t,T}^{N,i}) - \mathbf{E}[g(X_{t,T})|\mathcal{F}_t]\right) \Rightarrow \epsilon\tilde{K}_{t,T}(X_t) + M_{t,T}(X_t)$$

$$\sqrt{M}\left(\frac{1}{M}\sum_{i=1}^{M} g(\hat{X}_{t,T}^{N,i}) - \mathbf{E}[g(X_{t,T})|\mathcal{F}_t]\right) \Rightarrow \epsilon\hat{K}_{t,T}(X_t) + M_{t,T}(X_t)$$

respectively, for the scheme without and with transformation, where $\epsilon = \lim_{M,N\to\infty} \frac{\sqrt{M}}{N}$ *and*

$$[M_{t,\cdot}, M_{t,\cdot}]_T = \int_t^T \mathbf{E}[N_{t,s}(N_{t,s})'|\mathcal{F}_t]ds$$

$$N_{t,s} = \mathbf{E}[\partial g(X_{t,T})\mathcal{D}_s X_{s,T}|\mathcal{F}_s]$$

Remark 2 *(i) Assumption A is satisfied if* $d = 1$.

(ii) The asymptotic laws of the estimators have two parts. The first, K*, corresponds to the discretization bias; the second,* M*, results from the Monte Carlo estimation of the expectation. Note that* M *would not vanish even if we could sample from the law of* $X_{t,T}$*, since we are unable to calculate the conditional explication in closed form.*

(iii) It is clear from the results above that the transformation does not increase the speed of convergence of estimates of the expected value of a function of the terminal value of an SDE.

(iv) Both procedures have an asymptotic second order discretization bias given by $\epsilon\tilde{K}$ *and* $\epsilon\hat{K}$*, respectively. It follows that any confidence interval based on the Gaussian process* M *alone would suffer from a size distortion. Asymptotically valid confidence intervals must correct for this second order bias. Since the second order discretization bias was expressed as the expectation of random variables that can be simulated it is easy to correct for the size distortion, as we will show.*

(v) Both of our procedures achieve the same rate of convergence as the Euler scheme of an ODE. Additionally, any higher order scheme for the transformed state variables would not improve the speed of convergence of the overall procedure since speed is determined by the central limit theorem of independent replications involved in the estimation of the conditional expectation.

(vi) The second order discretization bias could be important in finite samples. Our explicit formulas can be used to assess its magnitude.

Before applying these results to the hedging term estimators we provide estimators that are asymptotically equivalent to standard Monte-Carlo estimators.

Let

$$\tilde{g}_{ct,T}^{N,M} = \frac{1}{M}\sum_{i=1}^{M}\left[g(\tilde{X}_{t,t+N\Delta}^{N,i}) + \partial g(\tilde{X}_{t,t+N\Delta}^{N,i})\tilde{C}_{t,t+N\Delta}^{N,i}\right]$$

with $n = 0, \dots, N-1$ and

$$\begin{aligned}
\tilde{C}_{t,t+(n+1)\Delta}^{N,i} &= \tilde{C}_{t,t+n\Delta}^{N,i} + [\partial A(\tilde{X}_{t,t+n\Delta}^{N,i})\Delta \\
&\quad + \sum_{j=1}^{d}\partial B_j(\tilde{X}_{t,t+n\Delta}^{N,i})(W_{t+(n+1)\Delta}^{j,i} - W_{t+n\Delta}^{j,i})]\tilde{C}_{t,t+n\Delta}^{N,i} \\
&\quad -[(\partial A)A](\tilde{X}_{t,t+n\Delta}^{N,i})\frac{\Delta}{N} \\
&\quad -\sum_{j=1}^{d}[(\partial A)B_j + (\partial B_j)A](\tilde{X}_{t,t+n\Delta}^{N,i})\frac{W_{t+(n+1)\Delta}^{j,i} - W_{t+n\Delta}^{j,i}}{N}
\end{aligned}$$

$$\tilde{X}_{t,t+(n+1)\Delta} - \tilde{X}_{t,t+n\Delta} = A(\tilde{X}_{t,t+n\Delta}^{N,i})\Delta + \sum_{j=1}^{d} B_j(\tilde{X}_{t,t+n\Delta}^{N,i})(W_{t+(n+1)\Delta}^{j,i} - W_{t+n\Delta}^{j,i})$$

with $\tilde{X}_{t,t}^{N,i} = X_t$ and $\tilde{C}_{t,t} = 0$ where $\Delta = \frac{T-t}{N}$.

Similarly if we apply first Doss transformation

$$\hat{g}_{ct,T}^{N,M} = \frac{1}{M}\sum_{i=1}^{M}\left[g(\hat{X}_{t,t+N\Delta}^{N,i}) + \partial g(\hat{X}_{t,t+N\Delta}^{N,i})\hat{C}_{t,t+N\Delta}^{N,i}\right]$$

with $n = 0, \dots, N-1$ and

$$\begin{aligned}
\hat{C}_{t,t+(n+1)\Delta}^{N,i} &= \tilde{C}_{t,t+n\Delta}^{N,i} + [\partial\hat{\mu}(\hat{X}_{t,t+n\Delta}^{N,i})\Delta]\hat{C}_{t,t+n\Delta}^{N,i} - [(\partial\hat{\mu})\hat{\mu}](\hat{X}_{t,t+n\Delta}^{N,i})\frac{\Delta}{N} \\
&\quad -\sum_{j=1}^{d}(\partial\hat{\mu})\frac{W_{t+(n+1)\Delta}^{i} - W_{t+n\Delta}^{i}}{N}
\end{aligned}$$

$$\hat{X}_{t,t+(n+1)\Delta} - \hat{X}_{t,t+n\Delta} = \hat{\mu}(\hat{X}_{t,t+n\Delta}^{N,i})\Delta + (W_{t+(n+1)\Delta}^{i} - W_{t+n\Delta}^{i})$$

with $\hat{X}_{t,t}^{N,i} = X_t$ and $\hat{C}_{t,t} = 0$ where $\Delta = \frac{T-t}{N}$.

As shown in the next theorem the estimators $\hat{g}^{N,M}$ and $\tilde{g}^{N,M}$ are asymptotically equivalent. They also have the same asymptotic distribution as the Monte Carlo estimator obtained by sampling from the true distribution of $X_{t,T}$.

Theorem 49 : *Under the conditions of theorem 48 we have*

$$\sqrt{M}(\tilde{g}_{ct,T}^{N,M} - \mathbf{E}[g(X_{t,T})|\mathcal{F}_t]) \Rightarrow M_{t,T}(X_t)$$

and

$$\sqrt{M}(\tilde{g}_{ct,T}^{N,M} - \hat{g}_{ct,T}^{N,M}) \Rightarrow 0$$

Remark 3 *(i) Standard asymptotic confidence intervals based on the Gaussian random variate do not any longer suffer from size distortion if we use the second order bias corrected estimators* $\tilde{g}^{N,M}$ *and* $\hat{g}^{N,M}$

(ii) Since the second order bias corrected estimator using the Doss transformation is asymptotically equivalent to the second order bias corrected estimator without, all potential advantages of the transformation have disappeared. In finite samples, however, these estimators may perform differently.

(iii) The asymptotic equivalence with an infeasible estimator based on sampling from the true distribution shows that the precision of the corrected estimators depends on the quadratic variation $[M_{t,\cdot}, M_{t,\cdot}]_T$. *When this variance is large we can use standard variance reduction techniques to obtain the shortest confidence intervals possible. It is important to note that variance reduction without second order bias correction may not be appropriate. Indeed, when the second order bias is large one may incorrectly conclude that precision is high since confidence intervals are short, when in fact this is due to the far smaller effective size of the confidence interval relative to its nominal size. This problem still exists if the approximation is performed with weak higher order schemes such as those proposed by Mil'shtein (1993). Since our second order bias corrected estimators are numerically not much more costly than those schemes there is a direct advantage to our procedure.*

3.6.2 Asymptotic Law of Hedging Term Estimators without Transformation.

Let us now apply the results above to our setting. Without transformation we estimate the functions a and b by

$$\tilde{a}_{k,t,T}^{N,M} = \frac{\sum_{i=1}^{M} \tilde{H}_{2,k,t,T}^{N,i}}{\sum_{i=1}^{M} \tilde{H}_{1,t,T}^{N,i}} \qquad \text{and} \qquad \tilde{b}_{k,t,T}^{N,M} = \frac{\sum_{i=1}^{M} \tilde{H}_{3,k,t,T}^{N,i}}{\sum_{i=1}^{M} \tilde{H}_{1,t,T}^{N,i}}$$

with $(\tilde{H}_{1,t,T}^{N,i}, \tilde{H}_{2,k,t,T}^{N,i}, \tilde{H}_{3,k,t,T}^{N,i}) = \tilde{f}(\tilde{X}_{k,t,T}^{N,i})$ as described before. The asymptotic distribution of these estimators is,

Theorem 50 *Under the conditions of theorem 46*

$$\sqrt{M}(\tilde{a}_{t,T}^{N,M} - a(t, Y_t)) \Rightarrow \epsilon \frac{\tilde{K}_{t,T}^{H_2}(Y_t)}{\mathbf{E}[\xi_{t,T}^{\rho} | \mathcal{F}_t]} + M_{t,T}^{a}(Y_t)$$

$$\sqrt{M}(\tilde{b}_{t,T}^{N,M} - b(t, Y_t)) \Rightarrow \epsilon \frac{\tilde{K}_{t,T}^{H_3}(Y_t)}{\mathbf{E}[\xi_{t,T}^{\rho} | \mathcal{F}_t]} + M_{t,T}^{b}(Y_t)$$

where $\epsilon = \lim_{M,N\to\infty} \frac{\sqrt{M}}{N}$ *and*

$$[M_{t,\cdot}^{a}, M_{t,\cdot}^{a}]_T = \int_t^T \frac{\mathbf{E}[N_{t,s}^{a}(N_{t,s}^{a})' | \mathcal{F}_t]}{\mathbf{E}[\xi_{t,T}^{\rho} | \mathcal{F}_t]} ds$$

$$[M_{t,\cdot}^{b}, M_{t,\cdot}^{b}]_T = \int_t^T \frac{\mathbf{E}[N_{t,s}^{b}(N_{t,s}^{b})' | \mathcal{F}_t]}{\mathbf{E}[\xi_{t,T}^{\rho} | \mathcal{F}_t]} ds$$

with

$$N_{k,t,s}^{a} = \mathbf{E}[\mathcal{D}_{k,s} H_{2,k,t,T} | \mathcal{F}_s]$$

$$N_{k,t,s}^{b} = \mathbf{E}[\mathcal{D}_{k,s} H_{3,k,t,T} | \mathcal{F}_s]$$

3.6.3 Asymptotic Law of Hedging Term Estimators with Transformation.

With transformation we estimate the function a and b by

$$\hat{a}^{N,M}_{k,t,T} = \frac{\sum_{i=1}^{M} \hat{H}^{N,i}_{2,k,t,T}}{\sum_{i=1}^{M} \hat{H}^{N,i}_{1,t,T}} \qquad \text{and} \qquad \tilde{b}^{N,M}_{k,t,T} = \frac{\sum_{i=1}^{M} \hat{H}^{N,i}_{3,k,t,T}}{\sum_{i=1}^{M} \hat{H}^{N,i}_{1,t,T}}$$

with $(\hat{H}^{N,i}_{1,t,T}, \hat{H}^{N,i}_{2,k,t,T}, \hat{H}^{N,i}_{3,k,t,T}) = \hat{f}(\hat{X}^{N,i}_{k,t,T})$ as described before. The asymptotic distribution of these estimators is given by

Theorem 51 *Under the conditions of theorem 46*

$$\sqrt{M}(\hat{a}^{N,M}_{t,T} - a(t, Y_t)) \Rightarrow \epsilon \frac{\hat{K}^{H_2}_{t,T}(Y_t)}{\mathbf{E}[\xi^{\rho}_{t,T}|\mathcal{F}_t]} + M^a_{t,T}(Y_t)$$

$$\sqrt{M}(\hat{b}^{N,M}_{t,T} - b(t, Y_t)) \Rightarrow \epsilon \frac{\tilde{K}^{H_3}_{t,T}(Y_t)}{\mathbf{E}[\xi^{\rho}_{t,T}|\mathcal{F}_t]} + M^b_{t,T}(X_t)$$

where ϵ, M^a, M^b are given as without transformation.

We see that the difference between the estimator with and without transformation is exclusively in the second order approximation bias. As mentioned before in general it will be difficult to establish whether or not this error is smaller for one of the two solution schemes.

In the next section we show how to correct for the second order bias. The resulting estimators are asymptotically equivalent with the infeasible estimator obtained if we could sample from the distributions of $\tilde{X}_{k,t,T}$ respectively $\hat{X}_{k,t,T}$.

3.6.4 Second Order Bias Corrected Estimators for Hedging Terms

As shown before since we have explicit solutions of the second order bias $\tilde{K}^{H_{2,k}}$, $\tilde{K}^{H_{3,k}}$ respectively $\hat{K}^{H_{2,k}}$, $\hat{K}^{H_{3,k}}$ as solutions of linear SDE's we have an easy way to correct for the second order bias involved in the asymptotic law of the functions determining the hedging terms in the optimal portfolio demand.

$$\tilde{a}_{ck,t,T}^{N,M} = \frac{\sum_{i=1}^{M}[\tilde{f}_2(\tilde{X}_{k,t,t+N\Delta}^{N,i}) + \partial\tilde{f}_2(\tilde{X}_{k,t,t+\Delta N}^{N,i})\tilde{C}_{k,t,t+\Delta N}^{N,i}]}{\sum_{i=1}^{M}\tilde{H}_{1,t,t+N\Delta}^{N,i}}$$

$$\tilde{b}_{ck,t,T}^{N,M} = \frac{\sum_{i=1}^{M}[\tilde{f}_3(\tilde{X}_{k,t,t+N\Delta}^{N,i}) + \partial\tilde{f}_3(\tilde{X}_{k,t,t+\Delta N}^{N,i})\tilde{C}_{k,t,t+\Delta N}^{N,i}]}{\sum_{i=1}^{M}\tilde{H}_{1,t,t+N\Delta}^{N,i}}$$

with

$$\begin{aligned}
\tilde{C}_{k,t,t+(n+1)\Delta}^{N,i} &= \tilde{C}_{k,t,t+n\Delta}^{N,i} + [\partial\tilde{A}(\tilde{X}_{k,t,t+n\Delta}^{N,i})\Delta \\
&+ \sum_{j=1}^{d}\partial\tilde{B}_j(\tilde{X}_{k,t,t+n\Delta}^{N,i})(W_{t+(n+1)\Delta}^{j,i} - W_{t+n\Delta})]\tilde{C}_{k,t,t+n\Delta}^{N,i} \\
&-[(\partial\tilde{A})\tilde{A}](\tilde{X}_{k,t,t+n\Delta}^{N,i}) \\
&\frac{\Delta}{N} - \sum_{j=1}^{d}[(\partial\tilde{A})\tilde{B}_j + (\partial\tilde{B}_j)\tilde{A}](\tilde{X}_{k,t,t+n\Delta}^{N,i})\frac{W_{t+(n+1)\Delta}^{j,i} - W_{t+n\Delta}^{j,i}}{N}
\end{aligned}$$

whereas

$$\tilde{X}_{k,t,t+(n+1)\Delta} - \tilde{X}_{k,t,t+n\Delta} = \tilde{A}(\tilde{X}_{k,t,t+n\Delta}^{N,i})\Delta + \sum_{j=1}^{d}\tilde{B}_j(\tilde{X}_{k,t,t+n\Delta}^{N,i})(W_{t+(n+1)\Delta}^{j,i} - W_{t+n\Delta}^{j,i})$$

with $\tilde{X}_{k,t,t}^{N,i} = \tilde{X}_{k,t}$ and $\tilde{C}_{k,t,t} = 0$ where $\Delta = \frac{T-t}{N}$.

Similarly if we apply first Doss transformation

$$\hat{a}_{ck,t,T}^{N,M} = \frac{\sum_{i=1}^{M}[\hat{f}_2(\hat{X}_{k,t,t+N\Delta}^{N,i}) + \partial\hat{f}_2(\hat{X}_{k,t,t+\Delta N}^{N,i})\hat{C}_{k,t,t+\Delta N}^{N,i}]}{\sum_{i=1}\hat{H}_{1,t,t+N\Delta}^{N,i}}$$

$$\hat{b}_{ck,t,T}^{N,M} = \frac{\sum_{i=1}^{M}[\hat{f}_3(\hat{X}_{k,t,t+N\Delta}^{N,i}) + \partial\hat{f}_3(\hat{X}_{k,t,t+\Delta N}^{N,i})\hat{C}_{k,t,t+\Delta N}^{N,i}]}{\sum_{i=1}\hat{H}_{1,t,T}^{N,i}}$$

with

$$\begin{aligned}
\hat{C}_{k,t,t+(n+1)\Delta}^{N,i} &= \tilde{C}_{k,t,t+n\Delta}^{N,i} + [\partial\hat{A}(\hat{X}_{k,t,t+n\Delta}^{N,i})\Delta]\hat{C}_{k,t,t+n\Delta}^{N,i} \\
&-[(\partial\hat{A})\hat{A}](\hat{X}_{k,t,t+n\Delta}^{N,i})\frac{\Delta}{N} \\
&-\sum_{j=1}^{d}\partial\hat{A}(\hat{X}_{k,t,t+n\Delta})\hat{B}_j\frac{W_{t+(n+1)\Delta}^{j,i} - W_{t+n\Delta}^{j,i}}{N}
\end{aligned} \tag{3.29}$$

whereas

$$\hat{X}_{k,t,t+(n+1)\Delta} - \hat{X}_{k,t,t+n\Delta} = \hat{A}(\hat{X}^{N,i}_{k,t,t+n\Delta})\Delta + \sum_{j=1}^{d} \hat{B}_j(W^{j,i}_{t+(n+1)\Delta} - W^{j,i}_{t+n\Delta}) \quad (3.30)$$

with $\hat{X}^{N,i}_{k,t,t} = \hat{X}_{k,t}$ and $\hat{C}_{t,t} = 0$ where $\Delta = \frac{T-t}{N}$.

We now get the following equivalent to the general theorem

Theorem 52 *Under the conditions of theorem 50 and 51 we have that*

$$\sqrt{M}(\tilde{a}^{N,M}_{ct,} - a(t, Y_t)) \Rightarrow M^a_{t,T}(Y_t) \quad (3.31)$$

$$\sqrt{M}(\tilde{b}^{N,M}_{ct,} - b(t, Y_t)) \Rightarrow M^b_{t,T}(Y_t). \quad (3.32)$$

Furthermore

$$\sqrt{M}(\tilde{a}^{N,M}_{ct,T} - \hat{a}^{N,M}_{ct,T}) \Rightarrow 0 \quad (3.33)$$

$$\sqrt{M}(\tilde{b}^{N,M}_{ct,T} - \hat{b}^{N,M}_{ct,T}) \Rightarrow 0 \quad (3.34)$$

3.7 Asymptotic Laws of Estimators with Unknown Initial State

Up to now we have assumed that the initial position of state variables X_t is known. In this section we extend our results and allow for situations where this is not any longer the case and future state variables have to be approximated. Such situations occur when we are interested in properties of future estimators from today's point of view. If we want to answer questions like what is the probability that the future hedging demand will be in a prescribed set.

We obtain such estimators in a straightforward way. Suppose we have information up to time t today, we then first simulate a trajectory up to time $t+\tau$. The corresponding feasible predictor of the future state is then given by $X^N_{t,t+\tau}(X_t)$. This value serves as starting point for M trajectories used to calculate estimators like $a_{t+\tau,T}$ respectively $b_{t+\tau,T}$ for example. If we want to analyze the asymptotic properties of approximation errors for such estimators we can again distinguish two sources of errors, a discretization error and Monte Carlo error. But now the need to discretize the SDE enters in two forms. First as before we cannot sample from the true random variables involved in the expectations of the functions we want to estimate and secondly now also the future state where we want to estimate can only be approximated.

As we will show now the difference in speed of convergence for the expected approximation error and the asymptotic distribution of the solution of an SDE without the transformation implies that only the latter approximation error is asymptotically not negligible. Furthermore in this case the speed of convergence of the overall procedure is slower. As a consequence estimators that are second order bias corrected are asymptotically equivalent to non-corrected estimators if initial state variables have also to be approximated.

On the other hand with transformation, the asymptotic distribution for second order bias corrected estimators is different than the one obtained without bias correction. In this case both discretization errors, the error coming from the approximation of the initial state and the discretization error of the approximated variables in the estimation of the conditional expectation are of the same order. Consequently the speed of the overall procedure is the same as if the initial state would have been known.

We first give results for the general SDE

$$dX_t = A(X_t)dt + \sum_{j=1}^{d} B_j(X_t)dW_t^j$$

and estimators

$$g^{N,M}_{t+\tau,T}(X^N_{t+\tau}) = \frac{1}{M}\sum_{i=1}^{d} g(X^N_{t+\tau,T}(X^N_{t+\tau}))$$

where $X_{t+\tau,T}(x)$ denotes the stochastic flow with starting point x at $t+\tau$.

Theorem 53 *Under the conditions of theorem 50 we have that*

$$\sqrt{M}(\tilde{g}^{N,M}_{t+\tau,T}(\tilde{X}^N_{t,t+\tau}) - \mathbf{E}[g(X_{t+\tau,T})|\mathcal{F}_{t+\tau}]) \;\Rightarrow\; M_{t+\tau,T}(X_{t,t+\tau}) + \varepsilon\mathbf{E}[\partial g(X_{t+\tau,T})\tilde{\Omega}_{t+\tau,T}|\mathcal{F}_t]\tilde{U}^X_{t,t+\tau} \tag{3.35}$$

where $\varepsilon = \lim_{M,N\to\infty}\frac{\sqrt{M}}{\sqrt{N}}$ *and* M *is given as in 48, respectively* $\tilde{\Omega}$ *as in 36. Furthermore we have*

$$\sqrt{M}(\tilde{g}^{N,M}_{ct+\tau,T}(X^N_{t,t+\tau}) - \tilde{g}^{N,M}_{t+\tau,T}(\tilde{X}^N_{t,t+\tau})) \Rightarrow 0 \tag{3.36}$$

Remark 4 1. *The speed of convergence is* $\frac{1}{\sqrt{N}}$. *If we double the number of Monte Carlo replications we must at least also double the number of discretization points to guarantee convergence. This is in contrast to estimators without transformation with known initial value. In this case the speed of convergence is* $\frac{1}{N}$ *and we have to double the number of discretization points only if the number of Monte Carlo replications is quadrupled. It follows that given a budget of computing time the estimators without transformation is more costly if the initial state is unknown.*

2. *The second order bias of the estimator with known initial value* $\tilde{K}$ *is asymptotically negligible since the asymptotic distribution of the state variables converges at a smaller speed. This explains why second order bias correction does not pay off in this situation.*

We know give corresponding results if we can use the transformation.

Theorem 54 *Under the conditions of theorem 51 we have that*

$$\sqrt{M}(\hat{g}^{N,M}_{t+\tau,T}(\hat{X}^N_{t,t+\tau}) - \mathbf{E}[g(X_{t+\tau,T}|\mathcal{F}_{t+\tau}]) \;\Rightarrow\; \epsilon\hat{K}_{t+\tau,T}(X_{t,t+\tau}) + M_{t+\tau,T}(X_{t,t+\tau}) + \epsilon\mathbf{E}[\partial g(X_{t+\tau,T})\hat{\Omega}_{t+\tau,T}|\mathcal{F}_t]\hat{U}^X_{t,t+\tau} \tag{3.37}$$

where $\epsilon = \lim_{M,N\to\infty}\frac{\sqrt{M}}{N}$ *as in theorem 51 and M is given as in 48, respectively* $\hat{\Omega}$ *as in 40. Furthermore we have*

$$\sqrt{M}(\hat{g}c^{N,M}_{t+\tau,T}(\hat{X}^N_{t+\tau}) - \mathbf{E}[g(X_{t+\tau,T}(X_{t,t+\tau}))|\mathcal{F}_{t+\tau}]) \;\Rightarrow\; M_{t+\tau,T}(X_{t,t+\tau}) + \epsilon\mathbf{E}[\partial g(X_{t+\tau,T})|\mathcal{F}_t]\hat{U}^X_{t,t+\tau} \tag{3.38}$$

Remark 5

1. *The speed of convergence of this estimator is* $\frac{1}{N}$ *as if the initial state would have been known. The increase in speed of convergence compared to the estimator without transformation explains why the second order bias term of the conditional estimators does not vanish asymptotically. If we quadruple the number of Monte Carlo replications we need only to double the number of discretization points to guarantee the same precision as for the estimator without transformation.*

2. *The asymptotic law of the estimator with transformation is non-centered indicating a second order bias. This second order bias does not vanish when we use the second order bias corrected conditional estimator. But since the speed of convergence is not the same, a comparison relative to second order bias is difficult.*

We will now specialize these results to derive the asymptotic law of portfolio estimators for market timing.

For this we need the derivatives of the functionals involved in the calculation of the hedging terms with respect to the initial state. We will use the following definition

Definition 55 *The derivative of the Brownian functionals $H_{j,t+\tau,T}$ with respect to the initial position $Y_{t+\tau,T}$ is as follows*

$$L^{H_2}_{t+\tau,T} = J_{1,t+\tau,T}J_{2,t+\tau,T}L^{J_0}_{t+\tau,T} + J_{0,t+\tau,T}J_{2,t+\tau,T}L^{J_1}_{t+\tau,T} + J_{0,t+\tau,T}J_{1,t+\tau,T}L^{J_2}_{t+\tau,T}$$

$$L^{H_3}_{t+\tau,T} = J_{1,t,T}J_{3,t+\tau,T}L^{J_1}_{t+\tau,T} + J_{0,t+\tau,T}J_{3,t+\tau,T}L^{J_1}_{t+\tau,T} + J_{0,t+\tau,T}J_{1,t+\tau,T}L^{J_3}_{t+\tau,T}$$

$$L^{J_0}_{t+\tau,T} = -\rho J_{0,t+\tau,T}\int_{t+\tau}^{T}\partial h(Y_s)\tilde{\Omega}_{3,t+\tau,s}ds$$

$$L^{J_1}_{t+\tau,T} = \rho J_{1,t+\tau,T}\int_{t+\tau}^{T}[[\sum_{j=1}^{d}\theta_j\partial\theta_j](Y_s)ds - \rho[\sum_{j=1}^{d}\partial\theta_j](Y_s)dW^j_s]\tilde{\Omega}_{3,t+\tau,s}$$

$$L^{J_2}_{t+\tau,T} = \int_{t+\tau}^{T}((\mathcal{D}_{t+\tau}Y_s)'\partial^2 r(Y_s)\tilde{\Omega}_{3,t+\tau,s})'ds + \int_{t+\tau}^{T}\partial r(Y_s)(L^{\mathcal{D}_{t+\tau}Y_s}_{t+\tau,s})'ds$$

$$\begin{aligned} L^{J_3}_{t+\tau,T} &= \int_{t+\tau}^{T}[\theta(Y_s)' + \rho(dW_s)'\partial\theta(Y_s)](L^{(\mathcal{D}_{t+\tau}Y_s')}_{t+\tau,s})' \\ &\quad \int_{t+\tau}^{T}[((\mathcal{D}_{t+\tau}Y_s)'\partial\theta(Y_s))'ds + \rho((\mathcal{D}_{t+\tau}Y_s)'\otimes(dW_s)')\partial^2\theta(Y_s)]\tilde{\Omega}_{3,t+\tau,s} \end{aligned} \tag{3.39}$$

with

$$\begin{aligned} L^{(\mathcal{D}_{t+\tau}Y')}_{t+\tau,T} &= (I\otimes\tilde{\Omega}_{2,t+\tau,T})\times \\ &\quad \int_{t+\tau}^{T}((\mathcal{D}_{t+\tau}Y_s)'\otimes\tilde{\Omega}^{-1}_{2,t+\tau,s})[\partial^2 A(Y_s)ds + \sum_{j=1}^{d}\partial^2 B_j(Y_s)dW^j_s]\tilde{\Omega}_{3,t+\tau,s} \\ &\quad -\sum_{j=1}^{d}\int_{t+\tau}^{T}((\mathcal{D}_{t+\tau}Y_s)'(\partial B_j(Y_s))'\otimes\tilde{\Omega}^{-1}_{2,t+\tau,s})\partial^2 B_j(Y_s)\tilde{\Omega}_{3,t+\tau,s}ds \end{aligned} \tag{3.40}$$

With these definitions we are now ready to present result for market timing estimators with and without transformation.

3.7.1 Asymptotic Laws for Market Timing: No Transformation

We have seen that it is sufficient to estimate functions a and b to obtain estimators of the hedging term. In this section we consider the following estimators.

$$\tilde{a}^{N,M}_{t+\tau,T}(\tilde{X}^N_{t+\tau}) = \frac{\sum_{i=1}^M \tilde{H}^{N,i}_{2,t+\tau,T}(\tilde{X}^N_{t+\tau})}{\sum_{i=1}^M \tilde{H}^{N,i}_{1,t+\tau,T}(\tilde{X}^N_{t+\tau})}$$

and

$$\tilde{b}^{N,M}_{t+\tau,T}(\tilde{X}^N_{t+\tau}) = \frac{\sum_{i=1}^M \tilde{H}^{N,i}_{3,t+\tau,T}(\tilde{X}^N_{t+\tau})}{\sum_{i=1}^M \tilde{H}^{N,i}_{1,t+\tau,T}(X^N_{t+\tau})}$$

Given the results in the previous section we obtain the following results:

Theorem 56 *Under the assumptions of theorem 50 we have that*

$$\sqrt{M}(\tilde{a}^{N,M}_{t+\tau,T}(\tilde{X}^N_{t,t+\tau}) - a(t+\tau, Y_{t+\tau})) \Rightarrow M^a_{t\tau,T}(Y_{t,t+\tau}) + \varepsilon\frac{\mathbf{E}[L^{H_2}_{t+\tau,T}|\mathcal{F}_{t+\tau}]\tilde{U}^{H_2}_{t,t+\tau}}{\mathbf{E}[\xi^{\rho}_{t+\tau,T}|\mathcal{F}_{t+\tau}]} \tag{3.41}$$

$$\sqrt{M}(\tilde{b}^{N,M}_{t+\tau,T}(\tilde{X}^N_{t,t+\tau}) - b(t+\tau, Y_{t+\tau})) \Rightarrow M^b_{t\tau,T}(Y_{t,t+\tau}) + \varepsilon\frac{\mathbf{E}[L^{H_3}_{t+\tau,T}|\mathcal{F}_{t+\tau}]\tilde{U}^{H_3}_{t,t+\tau}}{\mathbf{E}[\xi^{\rho}_{t+\tau,T}|\mathcal{F}_{t+\tau}]} \tag{3.42}$$

Furthermore,

$$\sqrt{M}(\tilde{a_c}^{N,M}(\tilde{X}^N_{t+\tau}) - \tilde{a_c}^{N,M}(\tilde{X}^N_{t+\tau})) \Rightarrow 0 \tag{3.43}$$

respectively

$$\sqrt{M}(\tilde{b_c}^{N,M}(\tilde{X}^N_{t+\tau}) - \tilde{b_c}^{N,M}(\tilde{X}^N_{t+\tau})) \Rightarrow 0 \tag{3.44}$$

where in all expressions $\varepsilon = \lim_{M,N\to\infty}\sqrt{\frac{M}{N}}$.

In the next section we will derive the same results with transformation and show that in contrast to the asymptotic law with known initial state, for market timing strategies the transformation increases the speed of convergence from $\frac{1}{\sqrt{N}}$ to $\frac{1}{N}$. Given the expressions of the asymptotic distributions the construction of confidence sets and test for market timing strategies is now straightforward.

3.7.2 Asymptotic Laws for Market Timing: Transformation

We have seen that it is sufficient to estimate functions a and b to obtain estimators of the hedging term. In this section we consider the following estimators.

$$\hat{a}^{N,M}_{t+\tau,T}(\hat{X}^N_{t+\tau}) = \frac{\sum_{i=1}^M \hat{H}^{N,i}_{2,t+\tau,T}(\hat{X}^N_{t+\tau})}{\sum_{i=1}^M \hat{H}^{N,i}_{1,t+\tau,T}(\hat{X}^N_{t+\tau})}$$

and

$$\hat{b}^{N,M}_{t+\tau,T}(\hat{X}^N_{t+\tau}) = \frac{\sum_{i=1}^M \hat{H}^{N,i}_{3,t+\tau,T}(\hat{X}^N_{t+\tau})}{\sum_{i=1}^M \hat{H}^{N,i}_{1,t+\tau,T}(X^N_{t+\tau})}$$

Given the results in the previous section we obtain the following results

Theorem 57 *Under the assumptions of theorem 51 we have that*

$$\sqrt{M}(\hat{a}^{N,M}_{t+\tau,T}(\hat{X}^N_{t,t+\tau}) - a(t+\tau, Y_{t+\tau})) \Rightarrow \epsilon\hat{K}^a_{t+\tau,T}(Y_{t,t+\tau}) + M^a_{t\tau,T}(Y_{t,t+\tau}) + \epsilon\frac{\mathbf{E}[L^{H_2}_{t+\tau,T}|\mathcal{F}_{t+\tau}]\hat{U}^{H_2}_{t,t+\tau}}{\mathbf{E}[\xi^\rho_{t+\tau,T}|\mathcal{F}_{t+\tau}]} \tag{3.45}$$

$$\sqrt{M}(\hat{b}^{N,M}_{t+\tau,T}(\hat{X}^N_{t,t+\tau}) - b(t+\tau, Y_{t+\tau})) \Rightarrow \epsilon\hat{K}^b_{t+\tau,T}(Y_{t,t+\tau}) + M^b_{t\tau,T}(Y_{t,t+\tau}) + \epsilon\frac{\mathbf{E}[L^{H_3}_{t+\tau,T}|\mathcal{F}_{t+\tau}]\hat{U}^{H_3}_{t,t+\tau}}{\mathbf{E}[\xi^\rho_{t+\tau,T}|\mathcal{F}_{t+\tau}]} \tag{3.46}$$

Furthermore,

$$\sqrt{M}(\hat{a}_c{}^{N,M}(\hat{X}^N_{t+\tau}) - \hat{a}_c{}^{N,M}(\hat{X}^N_{t+\tau} \Rightarrow -\epsilon\hat{K}^a_{t+\tau,T}(Y_{t,t+\tau}) \tag{3.47}$$

respectively

$$\sqrt{M}(\hat{b_c}^{N,M}(\hat{X}^N_{t+\tau}) - \hat{b_c}^{N,M}(\hat{X}^N_{t+\tau} \Rightarrow -\epsilon \hat{K}^b_{t+\tau,T}(Y_{t,t+\tau}) \tag{3.48}$$

where in all expressions $\epsilon = \lim_{M,N\to\infty} \frac{\sqrt{M}}{N}$ *and* $\hat{K}^a_{t+\tau,T}(Y_{t,t+\tau}) = \frac{\hat{K}^{H_2}_{t+\tau,T}(Y_{t,t+\tau})}{\mathbf{E}[\xi^\rho_{t+\tau,T}|\mathcal{F}_{t+\tau}]}$, *whereas* $\hat{K}^b_{t+\tau,T}(Y_{t,t+\tau}) = \frac{\hat{K}^{H_3}_{t+\tau,T}(Y_{t,t+\tau})}{\mathbf{E}[\xi^\rho_{t+\tau,T}|\mathcal{F}_{t+\tau}]}$.

3.8 Proofs

The next lemma is important to obtain the asymptotic law of the discretization errors.

Lemma 58 *The following weak convergence results hold*

$$V^{1,N}_t \equiv N \int_0^t (s - \eta^N_s) ds \Rightarrow \frac{1}{2} t \tag{3.49}$$

$$V^{2,i,N}_t \equiv N \int_0^t (W^i_s - W^i \circ \eta^N_s) ds \Rightarrow \frac{1}{2} W^i_t + \frac{1}{\sqrt{12}} B^i_t \tag{3.50}$$

$$V^{3,i,N}_t \equiv N \int_0^t (s - \eta^N_s) dW^i_s \Rightarrow \frac{1}{2} W^i_t - \frac{1}{\sqrt{12}} B^i_t \tag{3.51}$$

$$V^{4,i,j,N}_t \equiv \sqrt{N} \int_0^t (W^i_s - W^i \circ \eta^N_s) dW^j_s \Rightarrow \frac{1}{\sqrt{2}} B^{i,j}_t \tag{3.52}$$

where $((W^i)_{i\in\{1,\ldots,d\}}, (B^i)_{i\in\{1,\ldots,d\}}, (B^{i,j})_{i,j\in\{1,\ldots,d\}})$ *is a* $(2d+d^2)$*- dimensional standard Brownian motion.*

Proof of lemma 58: Since $\eta^N_t = \frac{[Nt]}{N}$ for $Nt \notin \mathbb{N}$ we have

$$\int_0^t (s - \eta^N_s) ds = \frac{1}{N^2} \int_0^{Nt} (s - [s]) ds.$$

Thus,

$$N\int_0^t (s-\eta_s^N)ds = \frac{1}{N}\sum_{k=1}^{[Nt]}\int_{[k-1,k[}(s-[s])ds + \frac{1}{N}\int_{[Nt]}^{Nt}(s-[s])ds$$

and therefore since $[s]=k-1$ for $s\in[k-1,k[$

$$N\int_0^t (s-\eta_s^N)ds = \frac{1}{N}\sum_{k=1}^{[Nt]}\int_0^1 s ds + \frac{1}{N}\int_{[Nt]}^{Nt}(s-[s])ds.$$

Result (3.49) then follows using

$$\frac{1}{N}\sum_{k=1}^{[Nt]}\int_0^1 s ds \to \frac{1}{2}t$$

and

$$\frac{1}{N}\int_{[Nt]}^{Nt}(s-[s])ds \to 0$$

when $N\to\infty$.

Similarly, we can show using the scaling property of Brownian motion that

$$N\int_0^t (W_s^i - W^i\circ\eta_s^N)ds = \frac{1}{\sqrt{N}}\sum_{k=1}^{[Nt]}\int_{[k-1,k[}(W_s^i - W_{[s]}^i)ds + \frac{1}{\sqrt{N}}\int_{[Nt]}^{Nt}(W_s^i - W_{[s]}^i)ds.$$

Itô's lemma then implies

$$\int_{[k-1,k[}(W_s^i - W_{[s]}^i)ds = \int_{k-1}^{k}(k-s)dW_s^i$$

so that

$$N\int_0^t (W_s^i - W^i\circ\eta_s^N)ds = \frac{1}{\sqrt{N}}\sum_{k=1}^{[Nt]}\int_{[k-1,k[}(k-s)dW_s^i + \frac{1}{\sqrt{N}}\int_{[Nt]}^{Nt}(W_s^i - W_{[s]}^i)ds.$$

Note that the sequence of i.i.d random variables $\int_{[k-1,k[}(k-s)dW_s^i$ has variance $\frac{1}{3}$ and covariance $\frac{1}{2}\delta_{i,j}$ with the Brownian motion W^j. It then follows from Donsker's functional central limit theorem that

$$\frac{1}{\sqrt{N}}\sum_{k=1}^{[Nt]}\int_{[k-1,k[}(k-s)dW_s^i \Rightarrow \frac{1}{2}W_t^i + \frac{1}{\sqrt{12}}B_t^i$$

where B^i is a standard Brownian motion independent of W^j for all $j \in \{1, \dots, d\}$. This establishes (3.50) since the continuity of the pathwise integral with respect to the Lebesgue measure implies that

$$\mathbf{P} - \lim \frac{1}{\sqrt{N}} \int_{[Nt]}^{Nt} (W_s^i - W_{[s]}^i) ds = 0.$$

The same type of argument establishes that

$$N \int_0^t (s - \eta_s^N) dW_s^i = \frac{1}{\sqrt{N}} \sum_{k=1}^{[Nt]} \int_{[k-1,k[} (s - [s]) dW_s^i + \frac{1}{\sqrt{N}} \int_{[Nt]}^{Nt} (s - [s]) dW_s^i.$$

Again, by Donsker's functional central limit theorem, the first part converges to a Brownian motion whereas the second part converges to zero in probability by the continuity of the Wiener integral

$$\mathbf{P} - \lim \frac{1}{\sqrt{N}} \int_{[Nt]}^{Nt} (s - [s]) dW_s^i = 0.$$

Since the sequence of i.i.d. random variables $\int_{[k-1,k[}(s - [s])dW_s^i$ has variance $\frac{1}{3}$ and covariance $\frac{1}{2}\delta_{i,j}$ with W^j as well as covariance $\frac{1}{6}\delta_{i,j}$ with $\int_{[k-1,k[}(k - s)dW_s^j$ we have

$$\frac{1}{\sqrt{N}} \sum_{k=1}^{[Nt]} \int_{[k-1,k[} (s - [s]) dW_s^i \Rightarrow \frac{1}{2} W_t^i - \frac{1}{\sqrt{12}} B_t^i.$$

This establishes (3.51).

It remains to show (3.52). Again by the scaling property of Brownian motion

$$\begin{aligned} \sqrt{N} \int_0^t (W_s^j - W^j \circ \eta_s^N) dW_s^i &= \frac{1}{\sqrt{N}} \sum_{k=1}^{[Nt]} \int_{[k-1,k[} (W_s^j - W_{[s]}^j) dW_s^i \\ &\quad + \frac{1}{\sqrt{N}} \int_{[Nt]}^{Nt} (W_s^j - W_{[s]}^j) dW_s^i. \end{aligned}$$

Since the sequence of i.i.d random variables $\int_{[k-1,k[}(W_s^j - W_{[s]}^j)dW_s^i$ has variance of $\frac{1}{2}$ and is independent of W^j, $\int_{[k-1,k[}(W_s^i - W_{[s]}^i)ds$ as well as of $\int_{[k-1,k[}(s - [s])dW_s^i$

we have again by Donsker's invariance principle that

$$\frac{1}{\sqrt{N}}\sum_{k=1}^{[Nt]}\int_{[k-1,k[}(W_s^j - W_{[s]}^j)dW_s^i \Rightarrow \frac{1}{\sqrt{2}}B_t^{i,j}.$$

whereas it follows from the continuity of the Itô integral that

$$\mathbf{P}-\lim \frac{1}{\sqrt{N}}\int_{[Nt]}^{Nt}(W_s^j - W_{[s]}^j)dW_s^i = 0.$$

This completes the proof of (3.52). ■

Duffie and Protter (1992) gives suitable conditions for the weak convergence of stochastic integrals. We adopt their definition of goodness.

Definition 59 (Duffie and Protter (1992)) *A sequence of semimartingales $\{X^N\}$ is good if, for any $\{H^N\}$ the convergence of (H^N, X^N) to (H, X) in distribution implies that X is a semimartingale and that $(H^N, X^N, \int H^N dX^N)$ converges to $(H, X, \int H_- dX)$ in distribution.*

Below we will see that the expansion of the approximation error involves stochastic integrals with respect to V^N. To establish the limit laws of our estimation errors we must prove that V^N is good.

Lemma 60 *The semimartingale V^N is good.*

Proof of Lemma 60: It follows from condition A of Duffie and Protter (1992) that it is sufficient for $V^{1,N}$ to be good that $N\int_0^T(s-\eta_s^N)ds < \infty$. But from the proof of the previous lemma we know that

$$N\int_0^T (s-\eta_s^N)ds = \frac{1}{2}\frac{[NT]}{N} + \frac{1}{N}\int_{[NT]}^{NT}(s-[s])ds$$

and therefore

$$N\int_0^T (s-\eta_s^N)ds < \frac{1}{2}(T-(T^2)-[T]^2)$$

which shows that $V^{1,N}$ is good. Similarly for $V^{2,i,N}$ we have to show that $N\mathbf{E}[\int_0^T |W_s^i - W^i \circ \eta_s^N|ds] < \infty$. But since

$$N\mathbf{E}[\int_0^T |W_s^i - W^i \circ \eta_s^N|ds] < N^2\int_0^T (s-\eta_s^N)ds$$

the same argument as for $V^{1,N}$ also shows that $V^{2,i,N}$ is also good.

Next, condition A of Duffie and Protter (1992) for martingales states that it is sufficient for goodness that their variance be bounded. Clearly

$$VAR[V_T^{3,i,N}] = N^2\int_0^T (s-\eta_s^N)^2 ds$$

$$VAR[V_T^{4,i,j,N}] = N^2\int_0^T (s-\eta_s^N)ds.$$

Thus, both semimartingales are good if

$$N^2\int_0^T (s-\eta_s^N)^2 ds < \infty.$$

But an argument similar to that for $V^{1,N}$ shows

$$N^2\int_0^T (s-\eta_s^N)^2 ds < \frac{1}{3}(T+T^3-[T]^3)$$

which establishes the goodness of $V^{3,i,N}$ and $V^{4,i,j,N}$. ■

Before proceeding to prove the main results of this article we clarify an important point. Up to now we have considered Euler continuous approximations given by

$$X_{t,T}^N = X_t + \int_t^T A(X_{t,\eta_s^N}^N)ds + \sum_{j=1}^d \int_t^T B(X_{t,\eta_s^N}^N)dW_s^j. \tag{3.53}$$

Unfortunately, these approximation are numerically infeasible since ..... But since we are only interested in the asymptotic law of the feasible Euler approximations

$$\tilde{X}^N_{t,v} = X^N_{t,\frac{[Nv]}{N}} \tag{3.54}$$

this is no problem since it follows by theorem 3.1. of Jacod and Protter (1998) that for locally Lipschitz continuous coefficients with linear growth the resulting approximation errors are asymptotically equivalent.

3.8.1 Asymptotic Laws of Malliavin Derivatives

Results without Transformation

The following expansion is crucial for the proofs of the theorems. The mean value theorem enables us to write the following expansion of the error $U^{\tilde{X}^N_{k,t}} = \tilde{X}^N_{k,t} - \tilde{X}_{k,t}$ of the Euler continuous approximation,

$$\begin{aligned} U^{\tilde{X}^N_k}_{t,T} &= \int_t^T \sum_{h=1}^{2d+4} \partial_h \tilde{A}(\tilde{X}_{k,t,s} + \lambda_{1,h} U^{\tilde{X}^N_{h,k}}_{t,s}) U^{\tilde{X}^N_{h,k}}_{t,s} ds \\ &+ \int_t^T \sum_{l=1}^{2d+4} \sum_{j=1}^{d} \partial_l \tilde{B}_j(\tilde{X}_{k,t,s} + \lambda_{2,l} U^{\tilde{X}^N_{l,k}}_{t,s}) U^{\tilde{X}^N_{l,k,t}}_{t,s} dW^j_s \\ &- \int_t^T \sum_{l=1}^{2d+4} \partial_l \tilde{A}(\tilde{X}^N_{k,t,s} + \lambda_{3,l} \bar{\tilde{U}}^{\tilde{X}^N_{l,k}}_{t,s}) \bar{\tilde{U}}^{\tilde{X}^N_{l,k}}_{t,s} ds \\ &- \int_t^T \sum_{l=1}^{2d+4} \sum_{j=1}^{d} \partial_l \tilde{B}_j(\tilde{X}^N_{k,t,s} + \lambda_{4,l} \bar{\tilde{U}}^{\tilde{X}^N_{l,k}}_{t,s}) \bar{\tilde{U}}^{\tilde{X}^N_{l,k}}_{t,s} dW^j_s \end{aligned} \tag{3.55}$$

where $\lambda_{\cdot,l} \in]0,1[$ for all $l = 1, \ldots, 2d+4$, e_l is the I^{th} unit vector and $\bar{\tilde{U}}^{\tilde{X}^N_{l,k}}_{t,s} = \tilde{X}^N_{l,k,t,s} - \tilde{X}^N_{l,k,t,\eta^N_s}$. Since

$$\bar{\tilde{U}}^{\tilde{X}^N_{l,k}}_{t,s} = \tilde{A}_l(\tilde{X}^N_{k,t,\eta^N_s})(s - \eta^N_s) + \sum_{h=1}^{d} \tilde{B}_{l,h}(\tilde{X}_{k,t,s})(W^h_s - W^h_{\eta^N_s}) \tag{3.56}$$

we obtain

$$
\begin{aligned}
U_{t,T}^{\tilde{X}_k^N} &= \int_t^T \sum_{l=1}^{2d+4} \partial_l \tilde{A}(\tilde{X}_{k,t,s} + \lambda_{1,l} e_l U_{t,s}^{\tilde{X}_{l,k}^N}) U_{t,s}^{\tilde{X}_{l,k}^N} ds \qquad (3.57)\\
&+ \int_t^T \sum_{l=1}^{2d+4} \sum_{j=1}^{d} \partial_l \tilde{B}_j(\tilde{X}_{k,t,s} + \lambda_{2,l} e_l U_{t,s}^{\tilde{X}_{l,k}^N}) U_{t,s}^{\tilde{X}_{l,k}^N} dW_s^j\\
&- \frac{1}{N} \int_t^T \sum_{l=1}^{2d+4} \partial_l \tilde{A}(\tilde{X}_{k,t,s}^N + \lambda_{3,l} e_l \bar{\tilde{U}}_{t,s}^{\tilde{X}_{l,k}^N}) \tilde{A}_l(\tilde{X}_{k,t,\eta_s^N}^N) dV_s^{1,N}\\
&- \frac{1}{N} \int_t^T \sum_{l=1}^{2d+4} \partial_l \tilde{A}(\tilde{X}_{k,t,s}^N + \lambda_{3,l} e_l \bar{\tilde{U}}_{t,s}^{\tilde{X}_{l,k}^N}) \sum_{j=1}^{d} \tilde{B}_{lj}(\tilde{X}_{k,t,\eta_s^N}^N) dV_s^{2,j,N}\\
&- \frac{1}{N} \int_t^T \sum_{l=1}^{2d+4} \sum_{j=1}^{d} \partial_l \tilde{B}_j(\tilde{X}_{k,t,s}^N + \lambda_{4,l} e_l \bar{\tilde{U}}_{t,s}^{\tilde{X}_{l,k}^N}) \tilde{A}_l(\tilde{X}_{k,t,\eta_s^N}^N) dV_s^{3,j,N}\\
&- \frac{1}{\sqrt{N}} \int_t^T \sum_{l=1}^{2d+4} \sum_{j=1}^{d} \partial_l \tilde{B}_j(\tilde{X}_{k,t,s}^N + \lambda_{4,l} e_l \bar{\tilde{U}}_{t,s}^{\tilde{X}_{l,k}^N}) \sum_{h=1}^{d} \tilde{B}_{lh}(\tilde{X}_{k,t,\eta_s^N}^N) dV_s^{4,h,j,N}.
\end{aligned}
$$

Proof of Theorem 36: Since $\lim_{N\to\infty} \eta_s^N = s$ and

$$\mathbf{P} - \lim_{N\to\infty} \bar{\tilde{U}}^{\tilde{X}_{l,k}^N} = \mathbf{P} - \lim_{N\to\infty} \tilde{U}^{\tilde{X}_{l,k,t}^N} = 0 \qquad (3.58)$$

and since, by lemma 60, V^N is good it follows that

$$\sqrt{N}\tilde{U}^{\tilde{X}^N} \Rightarrow \tilde{U}^{\tilde{X}} \qquad (3.59)$$

where

$$
\begin{aligned}
\tilde{U}_{t,T}^{\tilde{X}_k} &= \int_t^T \sum_{l=1}^{2d+4} \partial_l \tilde{A}(\tilde{X}_{k,t,s}) U_{t,s}^{\tilde{X}_{l,k}^N} ds \qquad (3.60)\\
&+ \int_t^T \sum_{l=1}^{2d+4} \sum_{j=1}^{d} \partial_l \tilde{B}_j(\tilde{X}_{k,t,s}) U_{t,s}^{\tilde{X}_{l,k}^N} dW_s^j\\
&- \frac{1}{\sqrt{2}} \int_t^T \sum_{l=1}^{2d+4} \sum_{j=1}^{d} \partial_l \tilde{B}_j(\tilde{X}_{k,t,s}^N) \sum_{h=1}^{d} \tilde{B}_{lh}(\tilde{X}_{k,t,s}) dB_s^{h,j}.
\end{aligned}
$$

The solution of this linear SDE corresponds to the result announced. ■

Proof of Corollary 38: For both approximations the result are obtained by simple calculus using the definitions of the functions $\tilde{A}, \tilde{B}$, the linear SDEs 3.60 and the mean value theorem:

$$\tilde{f}(\tilde{X}^N_{k,t,T}) - \tilde{f}(\tilde{X}_{k,t,T}) = \sum_{l=1}^{2d+4} \partial_l \tilde{f}(\tilde{X}_{k,t,T} + \lambda_l e_l U^{\tilde{X}^N_{l,k}}_{t,T}) U^{\tilde{X}^N_{l,k}}_{t,T} \tag{3.61}$$

such that

$$\sqrt{N} \tilde{f}(\tilde{X}^N_{k,t,T}) - \tilde{f}(\tilde{X}_{k,t,T})) \Rightarrow \partial \tilde{f}(\tilde{X}_{k,t,T}) \tilde{U}^{\tilde{X}_k}_{t,T}. \tag{3.62}$$

Results with Doss Transformation

We have seen that to find the asymptotic law without distribution the martingale part of the approximation is responsible for the order of convergence $\frac{1}{\sqrt{N}}$. An error free approximation of the martingale part of the SDE would obviously improve the order of convergence. In general this is infeasible since we do not know the law of the increments $\int_t^{t\frac{T}{N}} \sigma^Y_s dW_s$, except when the volatility coefficient is deterministic. But if assumption A is satisfied there exists a transformation of the state variables which has Gaussian increments that can be approximated without error.

If we apply the Doss transformation and sample Gaussian increments $(W^j_{t+(n+1)\frac{T-t}{N}} - W^j_{t+n\frac{T-t}{N}})$ then we obtain the following expansion of the error $U^{\hat{X}^N_{k,t}} = \hat{X}^N_{k,t} - \hat{X}_{k,t}$ of the Euler continuous approximation can without transformation be expanded as follows

$$\begin{aligned} U^{\hat{X}^N_{k,t}}_{t,T} &= \int_t^T \sum_{h=1}^{2d+4} \partial_h \hat{A}(\hat{X}_{k,t,s} + \lambda_{1,h} e_h U^{\hat{X}^N_{h,k}}_{t,s}) U^{\hat{X}^N_{h,k,t}}_{t,s} ds \\ &\quad - \int_t^T \sum_{h=1}^{2d+4} \partial_h \hat{A}(\hat{X}^N_{k,t,s} + \lambda_{3,h} e_h \bar{\tilde{U}}^{\hat{X}^N_{h,k,t}}_{t,s}) \bar{\tilde{U}}^{\hat{X}^N_{h,k,t}}_{t,s} ds \end{aligned} \tag{3.63}$$

where $\lambda_{\cdot,h} \in]0,1[$ for all $h = 1, \ldots, 2d+4$, e_h is the h^{th} unit vector and $\bar{\hat{U}}_{t,s}^{\hat{X}^N_{h,k,t}} = \hat{X}^N_{h,k,t,s} - \hat{X}^N_{h,k,t,\eta^N_s}$. Since

$$\bar{\hat{U}}_{t,s}^{\hat{X}^N_{h,k,t}} = \hat{A}_h(\hat{X}_{k,t,\eta^N_s})(s - \eta^N_s) + \sum_{j=1}^{d} \hat{B}_{h,j}(W^j_s - W^j_{\eta^N_s}) \tag{3.64}$$

we obtain

$$\begin{aligned} U_{t,T}^{\hat{X}^N_{k,t}} &= \int_t^T \sum_{h=1}^{2d+4} \partial_h \hat{A}(\hat{X}_{k,t,s} + \lambda_{1,h} e_h U_{t,s}^{\hat{X}^N_{h,k}}) U_{t,s}^{\hat{X}^N_{h,k}} ds \\ &\quad -\frac{1}{N}\int_t^T \sum_{h=1}^{2d+4} \partial_h \hat{A}(\hat{X}^N_{k,t,s} + \lambda_{3,h} e_h \bar{\hat{U}}_{t,s}^{\hat{X}^N_{h,k,t}}) \hat{A}_h(\hat{X}_{k,t,\eta^N_s}) dV_s^{1,N} \\ &\quad -\frac{1}{N}\int_t^T \sum_{h=1}^{2d+4} \partial_h \hat{A}(\hat{X}^N_{k,t,s} + \lambda_{3,h} e_h \bar{\hat{U}}_{t,s}^{\hat{X}^N_{h,k}}) \sum_{j=1}^{d} \hat{B}_{hj} dV_s^{2,j,N} \end{aligned} \tag{3.65}$$

Proof of Theorem 40: Since $\lim_{N\to\infty} \eta^N_s = s$ and

$$\mathbf{P} - \lim_{N\to\infty} \bar{\hat{U}}^{\hat{X}^N_{h,k}} = \mathbf{P} - \lim_{N\to\infty} \hat{U}^{\hat{X}^N_{h,k}} = 0 \tag{3.66}$$

and by lemma 60 V^N is good it follows that

$$N\hat{U}^{\hat{X}^N} \Rightarrow \hat{U}^{\hat{X}} \tag{3.67}$$

where

$$\begin{aligned} \hat{U}_{t,T}^{\hat{X}_k} &= \frac{1}{2}\int_t^T \sum_{h=1}^{2d+4} \partial_h \hat{A}(\hat{X}_{k,t,s}) U_{t,s}^{\hat{X}_{h,k}} ds \\ &\quad -\frac{1}{2}\int_t^T \sum_{h=1}^{2d+4} \partial_h \hat{A}(\hat{X}_{k,t,s}) \hat{A}_h(\hat{X}_{k,t,s}) ds \\ &\quad -\int_t^T \sum_{h=1}^{2d+4} \partial_h \hat{A}(\hat{X}_{k,t,s}) \sum_{j=1}^{d} \hat{B}_{hj}[\frac{1}{2}dW^j_s + \frac{1}{12}dB^j_s] \end{aligned} \tag{3.68}$$

This is SDE is linear and its solution corresponds to the result announced.

Proofs of Corollary 42 and Corollary 43: The result of both corollaries are obtained by simple calculus using the definitions of the functions $\hat{A}$ respectively $\hat{B}$ and the linear SDE's 3.68 and that again by the mean value theorem

$$\hat{f}(\hat{X}^N_{k,t,T}) - \hat{f}(\hat{X}_{k,t,T}) = \sum_{l=1}^{2d+4} \partial_l \hat{f}(\hat{X}_{k,t,T} + \lambda_l e_l U^{\hat{X}^N_{l,k}}_{t,T}) U^{\hat{X}^N_{l,k}}_{t,T} \tag{3.69}$$

such that

$$N(\hat{f}(\hat{X}^N_{k,t,T}) - \hat{f}(\hat{X}_{k,t,T})) \Rightarrow \partial \hat{F}(\hat{X}_{k,t,T}) \hat{U}^{\hat{X}_k}_{t,T}. \tag{3.70}$$

3.8.2 Expected Approximation Errors

In the previous section we have established the asymptotic laws of the approximation error and found the order of weak convergence. These results seem to contradict results summarized in Kloeden and Platen (1993) or Mil'shtein (1993) that state that the weak order of convergence of the Euler scheme is $\frac{1}{N}$. The apparent contradictions stems from the fact that they use different criterion to measure weak convergence. They define the weak order of convergence as order of convergence of the expected approximation error for a sufficient smooth class of functions (the class of Talay and Bally (1996) being the most general)[3]. Since convergence plus uniform integrability implies convergence of means we can use the asymptotic laws derived in the previous section to establish the expected approximation error. The following proofs can be seen as probabilistic counterparts

[3]The difference can be illustrated with the following example. Consider the estimator $\bar{Z}^N = \frac{1}{N}\sum_{i=1}^N Z^i$ where Z^i are i.i.d square integrable random variables. Kloeden and Platen et al. then consider the error $|\mathbf{E}[f(\bar{Z}^N)] - \mathbf{E}[f(Z)]| \leq \frac{\mathbf{VAR[Z]}}{N}$ Consequently they conclude that $\bar{Z}^N$ is of order $\frac{1}{N}$ (or order 1). In contrast since by the central limit theorem for i.i.d. random variates $\sqrt{N}(\bar{Z}^N - Z) \Rightarrow Z$ where $Z \sim N(0, \mathbf{VAR}[Z])$ we would say that $\bar{Z}^N$ is of order $\frac{1}{\sqrt{N}}$ (or of oder $\frac{1}{2}$). This explains the different conclusions

of the characterization of the approximation error in terms of a PDE by Talay and Tubaro (1991) and Bally and Talay (1996).

Since the asymptotic law of the approximation errors with Doss transformation is non-centered the proofs are straight forward in this case. On the other hand without transformation the asymptotic law has expectation zero. Therefore, some more work is needed to get expected approximation errors but as we show we can again derive results using weak convergence results for components of the approximation errors.

Expected Approximation Errors without Transformation

In this section we will proof the result for the system

$$d\tilde{X}_{k,t,s} = \tilde{A}(\tilde{X}_{k,t,s}) + \sum_{j=1}^{d} \tilde{B}_j(\tilde{X}_{k,t,s})dW_s^j \tag{3.71}$$

with $\tilde{X}_{k,t,t} = \tilde{X}_{k,t}$ known. The general results in section 3.4 are for and SDE with drift μ and volatility $\sigma_{\cdot j}$. The translation of results should be obvious.

We have seen in the proof of the asymptotic law of the approximation error that the order of convergence and the asymptotic law is crucially determined by the convergence of

$$\tilde{I}_{4,t,T}^N := \int_t^T \sum_{l=1}^{2d+4} \sum_{j=1}^{d} \partial_l \tilde{B}_j(\tilde{X}_{k,t,s}^N + \lambda_{4,l} e_l \bar{\tilde{U}}_{t,s}^{\tilde{X}_{l,k}^N}) \sum_{h=1}^{d} \tilde{B}_{lh}(\tilde{X}_{k,t,\eta_s^N}^N) dV_s^{4,h,j,N} \tag{3.72}$$

We will show now that $\sqrt{N}\tilde{I}_{4,t,T}^N$ is asymptotically negligible for expected approximation errors.

The following shows that there is a link between $V^{4,h,\cdot,N}$ and $V^{2,N}$

Lemma 61 *For $M_{t,v} = \sum_{j=1}^{d} \int_t^v \alpha_s^j dW_s$ where α is adapted and square integrable*

P *a.s. the following covariation result holds true*

$$\sqrt{N}[M_t, V_s^{4,h,j,N}]_v = \int_t^v \alpha_s^j dV_s^{2,h,N} \tag{3.73}$$

Proof of lemma 61: Since

$$d[W^i, V^{h,j,N}]_s = \sqrt{N}\mathbf{1}_{\{i=j\}}(W_s^h - W_{\eta_s^N}^h)ds \tag{3.74}$$

and $dV_s^{2,h,N} = N(W_s^h - W_{\eta_s^N}^h)ds$ the result announced follows.

This lemma will also implicitly be used for the following lemma

Lemma 62 *We have for random variables $G \in \mathbb{D}^{1,2}$ and and $d\times d$- dimensional adapted square integrable matrix process $\beta \in \mathbf{L}^2(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P}) \times [t,T]$ such that*

$$\lim_{r\to\infty} \sup_N \mathbf{E}[\mathbf{1}_{\{\|G\sqrt{N}\int_t^T \beta_s d\tilde{I}_{4,t,s}^N\| > r\}} G\sqrt{N} \int_t^T \beta_s d\tilde{I}_{4,t,s}^N | \mathcal{F}_t] = 0 \tag{3.75}$$

P- *a.s. that*

$$\sqrt{N}\mathbf{E}[G\int_t^T \beta_s d\tilde{I}_{4,t,s}^N)|\mathcal{F}_t] \to 0 \tag{3.76}$$

P- *a.s.*.

Proof of lemma 62: First it follows by the Clark-Ocone Formula that any $G \in \mathbb{D}^{1,2}$ can be written

$$G = \mathbf{E}[G|\mathcal{F}_t] + \sum_{j=1}^d \int_t^T \alpha_s^j dW_s^j \tag{3.77}$$

where $\alpha_s^j = \mathbf{E}[\mathcal{D}_{j,s}G|\mathcal{F}_s]$. It follows by lemma 61 that

$$\mathbf{E}[G\sqrt{N}\int_t^T \beta_s d\tilde{I}_{4,t,s}^N | \mathcal{F}_t] = \sum_{h,j=1}^d \mathbf{E}[\int_t^T \alpha_s^j \beta_s \gamma_{4,s}^{h,j,N} dV_s^{2,h,N} | \mathcal{F}_t] \tag{3.78}$$

where

$$\gamma_{4,s}^{h,j,N} = \sum_{l=1}^{2d+4} \partial_l \tilde{B}_j(\tilde{X}_{k,t,s}^N + \lambda_{4,l} e_l \bar{\bar{\tilde{U}}}_{t,s}^{\tilde{X}_{l,k}^N}) \tilde{B}_{lh}(\tilde{X}_{k,t,\eta_s^N}^N) \tag{3.79}$$

But then since

$$\sum_{h,j=1}^{d} \int_t^T \alpha_s^j \beta_s \gamma_{4,s}^{h,j,N} dV_s^{2,h,N} \Rightarrow V_{4,t,T} \tag{3.80}$$

where

$$V_{4,t,T} = \sum_{h,j=1}^{d} \int_t^T \alpha_s^j \beta_s [(\partial \tilde{B}_j) \tilde{B}_h](\tilde{X}_{k,t,s})[\frac{1}{2} dW_s^h + \frac{1}{\sqrt{12}} dB_s^h] \tag{3.81}$$

and by assumption (3.75) the LHS of (3.80) is uniformly integrable with respect the conditional probability measure we must have that

$$\sqrt{N}\mathbf{E}[G \int_t^T \beta_s dU_{4,t,s}^N | \mathcal{F}_t] \rightarrow \mathbf{E}[V_{4,t,T} | \mathcal{F}_t] \tag{3.82}$$

and the result follows since $\mathbf{E}[V_{4,t,T}|\mathcal{F}_t] = 0$. This establishes the result announced.

With this two lemmas we are no able to proof theorem 44

Proof of theorem 44: It follows from the error expansion (3.57) that the approximation error $U^{\tilde{X}_k^N}$ satisfies a linear SDE. The solution can therefore be written

$$\begin{aligned} N(\tilde{\Omega}_{t,T}^N)^{-1} U_{t,T}^{\tilde{X}_k^N} &= -\int_t^T (\tilde{\Omega}_{t,s}^N)^{-1} d\tilde{I}_{1,t,s}^N - \int_t^T (\tilde{\Omega}_{t,s}^N)^{-1} d\tilde{I}_{2,t,s}^N \\ &\quad - \int_t^T (\tilde{\Omega}_{t,s}^N)^{-1} d[\tilde{I}_{3,t,s}^N - [\tilde{R}^N, \tilde{I}_{3,t}^N]_s] \\ &\quad - \sqrt{N} \int_t^T (\tilde{\Omega}_{t,s}^N)^{-1} d[\tilde{I}_{4,t,s}^N - [\tilde{R}^N, \tilde{I}_{4,t}^N]_s] \end{aligned} \tag{3.83}$$

where

$$\begin{aligned} \tilde{\Omega}_{t,T}^N &= \exp\left(\int_t^T \sum_{l=1}^{2d+4} \partial_l \tilde{A}(\tilde{X}_{k,t,s} + \lambda_{1,l} e_l U_{t,s}^{\tilde{X}_{l,k}^N}) ds + \int_t^T \sum_{j=1}^{d} \partial_l \tilde{B}_j (\tilde{X}_{k,t,s} + \lambda_{2,l} e_l U_{t,s}^{\tilde{X}_{l,k}^N}) dW_s^j \right) \\ &\quad \times \exp\left(-\frac{1}{2} \int_t^T \sum_{j=1}^{d} \partial_l \tilde{B}_j (\tilde{X}_{k,t,s} + \lambda_{2,l} e_l U_{t,s}^{\tilde{X}_{l,k}^N}) (\sum_{j=1}^{d} \partial_l \tilde{B}_j (\tilde{X}_{k,t,s} + \lambda_{2,l} e_l U_{t,s}^{\tilde{X}_{l,k}^N})' ds \right) \end{aligned}$$

and

$$\tilde{I}^N_{1,t,T} \equiv \int_t^T \sum_{l=1}^{2d+4} \partial_l \tilde{A}(\tilde{X}^N_{k,t,s} + \lambda_{3,l} e_l \bar{\tilde{U}}^{\tilde{X}^N_{l,k}}_{t,s}) \tilde{A}_l(\tilde{X}^N_{k,t,\eta^N_s}) dV^{1,N}_s$$

$$\tilde{I}^N_{2,t,T} \equiv \int_t^T \sum_{l=1}^{2d+4} \partial_l \tilde{A}(\tilde{X}^N_{k,t,s} + \lambda_{3,l} e_l \bar{\tilde{U}}^{\tilde{X}^N_{l,k}}_{t,s}) \sum_{j=1}^{d} \tilde{B}_{lj}(\tilde{X}^N_{k,t,\eta^N_s}) dV^{2,j,N}_s$$

$$\tilde{I}^N_{3,t,T} \equiv \int_t^T \sum_{l=1}^{2d+4} \sum_{j=1}^{d} \partial_l \tilde{B}_j(\tilde{X}^N_{k,t,s} + \lambda_{4,l} e_l \bar{\tilde{U}}^{\tilde{X}^N_{l,k}}_{t,s}) \tilde{A}_l(\tilde{X}^N_{k,t,\eta^N_s}) dV^{3,j,N}_s$$

whereas $\tilde{I}^N_{4,t,T}$ is given in (3.72).

Then since by lemma 61

$$\sqrt{N} d[\tilde{R}^N, \tilde{I}^N_{4,t}]_s = \sum_{h,j=1}^{d} \partial \tilde{B}_j(\tilde{X}_{k,t,s} + \lambda_{2,l} e_l U^{\tilde{X}^N_{l,k}}_{t,s}) \gamma^{h,j,N}_{4,s} dV^{2,h,N}_s \tag{3.84}$$

and therefore

$$\sqrt{N}[\tilde{R}^N, \tilde{I}^N_{4,t}]_T \Rightarrow \int_t^T \sum_{h,j=1}^{d} [\tilde{B}_j \partial \tilde{B}_j \tilde{B}_h](\tilde{X}_{k,t,s})[\frac{1}{2} dW^j_s + \frac{1}{\sqrt{12}} dB^j_s]. \tag{3.85}$$

Similarly, since

$$d[\tilde{R}^N, \tilde{I}^N_{3,t}]_s = \sum_{j=1}^{d} \partial \tilde{B}_j(\tilde{X}_{k,t,s} + \lambda_{2,l} e_l U^{\tilde{X}^N_{l,k}}_{t,s}) \gamma^{j,N}_{3,s} dV^{1,N}_s \tag{3.86}$$

with

$$\gamma^{j,N}_{3,s} = \sum_{l=1}^{d} \partial_l \tilde{B}_j(\tilde{X}_{k,t,s} + \lambda_{2,l} e_l U^{\tilde{X}^N_{l,k}}_{t,s}) \tilde{A}_l(\tilde{X}^N_{k,t,\eta^N_s}) \tag{3.87}$$

we have that

$$[\tilde{R}^N, \tilde{I}^N_{3,t}]_T \Rightarrow \int_t^T \sum_{j=1}^{d} [\partial \tilde{B}_j \partial \tilde{B}_j \tilde{A}](\tilde{X}_{k,t,s}) ds \tag{3.88}$$

For the remaining terms we have the following weak convergence results

$$(\tilde{I}^{1,N}, \tilde{I}^{2,N}, \tilde{I}^{3,N}) \Rightarrow (\tilde{I}^1, \tilde{I}^2, \tilde{I}^3) \tag{3.89}$$

where

$$\tilde{I}^1_{t,T} = \frac{1}{2}\int_t^T \partial\tilde{A}(\tilde{X}_{k,t,s})\tilde{A}(\tilde{X}_{k,t,s})ds \tag{3.90}$$

$$\tilde{I}^2_{t,T} = \int_t^T \sum_{j=1}^d \partial\tilde{A}(\tilde{X}_{k,t,s})\tilde{B}_j(\tilde{X}_{k,t,s})[\frac{1}{2}dW^j_s + \frac{1}{\sqrt{12}}dB^j_s] \tag{3.91}$$

$$\tilde{I}^3_{t,T} = \int_t^T \sum_{j=1}^d \partial\tilde{B}_j(\tilde{X}_{k,t,s})\tilde{A}(\tilde{X}_{k,t,s})[\frac{1}{2}dW^j_s - \frac{1}{\sqrt{12}}dB^j_s] \tag{3.92}$$

We show now that the only term that is of smaller order of weak convergence convergence to zero in expectation. This can be seen as follows. By lemma 62 with $G = [\tilde{\Omega}^N_{t,T}]_{i,j}$ and $\beta_s = \tilde{\Omega}^N_{t,s}$ and by assumption 3.23 we have that

$$\sqrt{N}\mathbf{E}[\tilde{\Omega}^N_{t,T}\int_t^T (\tilde{\Omega}^N_{t,T})^{-1}d\tilde{I}^N_{4,t,s}|\mathcal{F}_t] \to 0 \tag{3.93}$$

P- a.s. as $N \to \infty$.

Then since under the assumption on the function g and again by the fact that $\tilde{X}^N$ is convergent it follows by the mean value theorem that

$$N\mathbf{E}[g(\tilde{X}^N_{k,t,T}) - g(\tilde{X}_{k,t,T})|\mathcal{F}_t] \to \frac{1}{2}\mathbf{E}[\partial g(\tilde{X}_{k,t,T})\tilde{V}_{t,T}|\mathcal{F}_t] \tag{3.94}$$

where $V_{t,T}$ is obtained from the limit of the error $NU^N_{t,T}$ by taking into account that by lemma 62

$$\sqrt{B}\mathbf{E}[\partial g(\tilde{X}_{k,t,T})\tilde{\Omega}^N_{t,T}\int_t^T (\tilde{\Omega}^N_{t,T})^{-1}d\tilde{I}^N_{4,t,s}|\mathcal{F}_t] \to 0 \tag{3.95}$$

P- a.s. when $N \to \infty$ and that expectations involving stochastic integrals with respect to the BM B^j that are independent from $\mathcal{F}_t$ vanish. More precisely

$$2V_{t,T} = -\tilde{\Omega}_{t,T}\int_t^T \tilde{\Omega}_{t,s}^{-1}[\partial\tilde{A}(\tilde{X}_{k,t,s})d\tilde{X}_{k,t,s} + \sum_{j=1}^{d}[\partial\tilde{B}_j\tilde{A}](\tilde{X}_{k,t,s})dW_s^j$$
$$-\sum_{j=1}^{d}[\partial\tilde{B}_j\partial\tilde{B}_j](\tilde{X}_{k,t,s})d\tilde{X}_{k,t,s}]$$

This establishes the result announced.

The proof of the result for the approximation error of the random variables used to calculate the hedging terms is now straightforward.

Proof of Corollary 46: If we put in the general result $g = \tilde{f}$ and use the definitions of $\tilde{A}$ and $\tilde{B}_j$ we obtain the result announced.

Expected Approximation Error with Transformation

We now prove results for the expected approximation error using the Doss transformation.

Proof of Theorem 45: In this case all terms of the error expansion (3.65 are of order $1/N$ and the limit distribution is non-centered. The result is then obtained, if the conditional measure is uniformly integrable, by taking the expectation of the solution of the linear SDE for the approximation error

$$NU_{t,T}^{\hat{X}_{k,t}^N} = -\hat{\Omega}_{t,T}\int_t^T (\hat{\Omega}_{t,s}^N)^{-1}\int_t^T (\hat{\Omega}_{t,s}^N)^{-1}d\hat{I}_{1,t,s}^N - \int_t^T (\hat{\Omega}_{t,s}^N)^{-1}d\hat{I}_{2,t,s}^N \qquad (3.96)$$

where $\hat{\Omega}_{t,T}^N = \exp\left(\int_t^T \partial\hat{A}(\hat{X}_{k,t,s} + \lambda_{1,l}e_l U_{t,s}^{\hat{X}_{l,k}^N})ds\right)$ and

$$\hat{I}_{1,t,T}^N \equiv \int_t^T \sum_{l=1}^{2d+4} \partial_l\hat{A}(\hat{X}_{k,t,s}^N + \lambda_{3,l}e_l\bar{\hat{U}}_{t,s}^{\hat{X}_{l,k}^N})\hat{A}_l(\hat{X}_{k,t,\eta_s^N}^N)dV_s^{1,N}$$

$$\hat{I}^N_{2,t,T} \equiv \int_t^T \sum_{l=1}^{2d+4} \partial_l \hat{A}(\hat{X}^N_{k,t,s} + \lambda_{3,l} e_l \bar{\hat{U}}^{\hat{X}^N_{l,k}}_{t,s}) \sum_{j=1}^{d} \hat{B}_{lj}(\hat{X}^N_{k,t,\eta^N_s}) dV^{2,j,N}_s.$$

Since

$$(\hat{R}^N_t, \hat{\Omega}^N_t, I^{1,N}_t, I^{2,N}_t, \hat{X}^N_{k,t}) \Rightarrow (\hat{R}_t, \hat{\Omega}_t, I^1_t, I^2_t, \hat{X}_{k,t}) \tag{3.97}$$

where $\hat{R}_t, \hat{\Omega}_t, \hat{X}_{k,t}$ are as announced and

$$\hat{I}^1_{t,T} = \frac{1}{2}\int_t^T [(\partial\hat{A})\hat{A}](\hat{X}_{k,t,s})ds \tag{3.98}$$

$$\hat{I}^2_{t,T} = \int_t^T [[\partial A \sum_{j=1}^{d} \hat{B}_j](\hat{X}_{k,t,s})[\frac{1}{2}dW^j_s + \frac{1}{\sqrt{12}}dB^j_s] \tag{3.99}$$

The proof of the theorem now follows using the same arguments as in the proof without transformation. ■

Proof of Corollary 47: Substituting $g = \tilde{f}$ in the general result and using the definitions of $\tilde{A}$ and $\tilde{B}_j$ yields the result announced. ■

3.8.3 Asymptotic Law of Hedging Terms

Proof of Theorem 48: The proof with and without transformation is the same. The approximation error can be written

$$\frac{1}{\sqrt{M}}\sum_{i=1}^{M} g(X^{i,N}_{t,T}) - \mathbf{E}[g(X_{t,T})|\mathcal{F}_t] = \frac{1}{\sqrt{M}}\sum_{i=1}^{M}(g(X^{i,N}) - g(X^i)) - \sum_{i=1}^{M}(g(X^i) - E[g(X)|\mathcal{F}_t])$$

By the Lindeberg central limit theorem for i.i.d. random variables we then have

$$\sum_{i=1}^{M}(g(X^i) - E[g(X)|\mathcal{F}_t]) \Rightarrow VAR[g(X_{t,T})|\mathcal{F}_t]Z$$

where $Z \sim N(0,1)$. Since by the Clark-Ocone formula we have that

$$g(X_{t,T}) - \mathbf{E}[g(X_{t,T})|\mathcal{F}_t] = \int_t^T \mathbf{E}[\mathcal{D}_s g(X_{t,T})|\mathcal{F}_s]dW_s$$

the conditional variance can be written

$$VAR[g(X_{t,T}|\mathcal{F}_t] = \int_t^T \mathbf{E}[\|\partial g(X_{t,T})\mathcal{D}_s X_{t,T}\|^2|\mathcal{F}_s]ds$$

This establishes that $VAR[g(X_{t,T})|\mathcal{F}_t]Z = M_{t,T}(X_t)$. It remains to find the weak limit of $\frac{1}{\sqrt{M}}\sum_{i=1}^{M}(g(X_{t,T}^{i,N}) - g(X_{t,T}^{i}))$. We introduce a sequence of numbers dependent on N such that $\epsilon^N N = M$. We then have by the Kolmogorov's strong law of large numbers that for all N

$$N\epsilon^N \left(\frac{1}{M}\sum_{i=1}^{N}(g(X_{t,T}^{i,N}) - g(X_{t,T}^{i}))\right) \to N\epsilon^N \mathbf{E}[g(X_{t,T}^{N}) - g(X_{t,T})|\mathcal{F}_t]\ \mathbf{P}\text{-a.s.}$$

Then if $\lim_{N\to\infty}\epsilon^N = \epsilon < \infty$. We obtain from the previous results on the expected approximation error that

$$\frac{1}{\sqrt{M}}\sum_{i=1}^{M}(g(X_{t,T}^{i,N}) - g(X_{t,T}^{i})) \to \epsilon K(X_t)$$

as $N \to \infty$

This establishes the result announced

■

Proof of theorem 50 and theorem 51

Again the proof with and without transformation is the same.

Given the fact that $\mathbf{P} - lim\frac{1}{M}\sum_{i=1}^{M} H_{1,t,T}^{i,t,T} = \mathbf{E}[\xi_{t,T}^{\rho}|\mathcal{F}_t]$ we have that

$$\sqrt{M}(a_{t,T}^{N,M} - a(t,Y_t)) = \frac{\frac{1}{\sqrt{M}}\sum_{i=1}^{M} H_{2,t,T}^{N,i} - \mathbf{E}[H_{2,t,T}|\mathcal{F}_t]}{\mathbf{E}[\xi_{t,T}^{\rho}|\mathcal{F}_t]} + o_P(1)$$

respectively

$$\sqrt{M}(b_{t,T}^{N,M} - b(t,Y_t)) = \frac{\frac{1}{\sqrt{M}}\sum_{i=1}^{M} H_{3,t,T}^{N,i} - \mathbf{E}[H_{3,t,T}|\mathcal{F}_t]}{\mathbf{E}[\xi_{t,T}^{\rho}|\mathcal{F}_t]} + o_P(1).$$

The results announced follow then by the same arguments used to proof the general case ■

Proof of Theorem 49 The Euler approximation of the process $N\tilde{C}^{N,i}_{t,\cdot}$ given by

$$\begin{aligned}\tilde{C}^{N,i}_{t,t+(n+1)\Delta} &= \tilde{C}^{N,i}_{t,t+n\Delta} + [\partial A(\tilde{X}^{N,i}_{t,t+n\Delta})\Delta \\ &+ \sum_{j=1}^{d} \partial B_j(\tilde{X}^{N,i}_{t,t+n\Delta})(W^{j,i}_{t+(n+1)\Delta} - W^{j,i}_{t+n\Delta})]\tilde{C}^{N,i}_{t,t+n\Delta} \\ &- [(\partial A)A](\tilde{X}^{N,i}_{t,t+n\Delta})\frac{\Delta}{N} \\ &- \sum_{j=1}^{d}[(\partial A)B_j + (\partial B_j)A](\tilde{X}^{N,i}_{t,t+n\Delta})\frac{W^{j,i}_{t+(n+1)\Delta} - W^{j,i}_{t+n\Delta}}{N}\end{aligned}$$

has a Euler continuous approximation

$$\begin{aligned}d(N\tilde{C}^{N}_{t,s}) &= [\partial A(\tilde{X}^{N}_{t,\eta^N_s})ds + \sum_{j=1}^{d} \partial B_j(\tilde{X}^{N}_{t,\eta^N_s})dW^j_s](N\tilde{C}^{N}_{t,\eta^N_s}) \\ &- [(\partial A)A](\tilde{X}^{N}_{t,\eta^N_s} ds - \sum_{j=1}^{d}[(\partial A)B_j + (\partial B_j)A](\tilde{X}_{t,\eta^N_s})dW^j_s\end{aligned}$$

It then follows that

$$N\mathbf{E}[\partial g(\tilde{X}^{N}_{t,N\Delta})\tilde{C}^{N}|\mathcal{F}_t] \to -\tilde{K}(X_t) \tag{3.100}$$

Consequently, by the same arguments as in the proof of theorem 48 we have for $\epsilon^N = \frac{\sqrt{M}}{N}$ such that $\lim_{N\to\infty}\epsilon^N = \epsilon < \infty$ that

$$\frac{1}{\sqrt{M}}\sum_{i=1}^{M} \partial g(\tilde{X}^{N,i}_{t,t+N\Delta})\tilde{C}^{N,i}_{t,t+N\Delta} \to -\epsilon^N N\mathbf{E}[\partial g(\tilde{X}^{N}_{t,T})\tilde{C}^{N}_{t,T}|\mathcal{F}_t]$$

as $M \to \infty$. It follows that the term $\frac{1}{\sqrt{M}}\sum_{i=1}^{M} \partial g(\tilde{X}^{N,i}_{t,t+N\Delta})\tilde{C}^{N,i}_{t,t+N\Delta}$ corrects the asymptotic second order bias for the estimator without transformation.

The proof of the result for the estimator with transformation follows exactly the same steps. In this case the average over independent replications of the random variables $\partial g(\hat{X}_{t,T}^N)\hat{C}_{t,t+n\Delta}^N$ approximate the negative of the second order bias with transformation.

The asymptotic equivalence of the bias corrected estimators with and without transformation is then simply a consequence of them having the same asymptotic distribution. ■

Proof of Theorem 52: The proof of this theorem follows again along the lines of the arguments used to establish the general result. The second order error correction variables $\tilde{C}^N$, respectively $\hat{C}^N$ are again chosen such that they approximate random variables whose expectation converges to the negative of the second order bias. Consequently, the estimators are asymptotically free of any effects from discretization. The asymptotic equivalence of the corrected estimators with and without transformation is simply a consequence of the fact that the part of the error distribution that captures the Monte Carlo approximation error is for both the same. ■

3.8.4 Asymptotic Laws of Estimators with Unkown Initial State

The estimators for market timing are obtained by first simulating trajectories up to time $t+\tau$ for $\tau \in [0, T-t[$ and then starting of M trajectories starting at $X_{t+\tau}^N$. Corresponding estimators are the given by

$$a_{t+\tau,T}^{N,M}(X_{t+\tau}^N) = \frac{\sum_{i=1}^M H_{2,t+\tau,T}^{N,i}(X_{t+\tau}^N)}{\sum_{i=1}^M H_{1,t+\tau,T}^{N,i}(X_{t+\tau}^N)}$$

and

$$b_{t+\tau,T}^{N,M}(X_{t+\tau}^N) = \frac{\sum_{i=1}^M H_{3,t+\tau,T}^{N,i}(X_{t+\tau}^N)}{\sum_{i=1}^M H_{1,t+\tau,T}^{N,i}(X_{t+\tau}^N)}$$

where we have made explicitly clear that random variables H^N are obtained with approximated starting point X^N. Again we obtain corresponding expressions with transformation for $X^N_{t+\tau} = \tilde{X}^N_{t+\tau}$ and $H^N_{\cdot,t,T}(X^N_{t+\tau}) = \tilde{H}^N_{\cdot,t+\tau,T}(\tilde{X}^N_{t+\tau})$ respectively estimators with transformation for $X^N_{t+\tau} = \hat{X}^N_{t+\tau}$ and $H^N_{\cdot,t+\tau,T}(X^N_{t+\tau}) = \hat{H}^N_{\cdot,t,T}(\hat{X}^N_{t+\tau})$.

We first consider a more general case and are interested to find asymptotic distributions of the errors $U^{g^{N,M}}_{t+\tau,T} = \sqrt{M}(g^{N,M}_{t+\tau,T}(X^N_{t+\tau}) - \mathbf{E}[g(X_{t+\tau,T}(X_{t+\tau}))|\mathcal{F}_{t+\tau}])$, that is with transformation

$$U^{\tilde{g}^{N,M}}_{t+\tau,T} = \frac{1}{\sqrt{M}}\sum_{i=1}^{M}(g(\tilde{X}^{N,i}_{t+\tau,T}(\tilde{X}^N_{t+\tau})) - \mathbf{E}[g(X_{t+\tau,T}(X_{t+\tau}))|\mathcal{F}_{t+\tau}])$$

and

$$U^{\hat{g}^{N,M}}_{t+\tau,T} = \frac{1}{\sqrt{M}}\sum_{i=1}^{M}(g(\hat{X}^{N,i}_{t+\tau,T}(\hat{X}^N_{t+\tau})) - \mathbf{E}[g(X_{t+\tau,T}(X_{t+\tau}))|\mathcal{F}_{t+\tau}])$$

without.

Note that in this case $X_{t+\tau,T}(X_{t+\tau})$ corresponds to the stochastic flow and therefore solves the SDE

$$X_{t+\tau,T}(X_{t+\tau}) = X_{t+\tau} + \int_{t+\tau}^{T} A(X_{t+\tau,s}(X_{t+\tau}))ds + \int_{t+\tau}^{T} B(X_{t+\tau,s}(X_{t+\tau}))dW_s$$

whereas $X^N_{t+\tau,T}(X^N_{t+\tau})$ corresponds to the numerical approximation of the stochastic flow. These will play an important role in the proofs of estimators for market timing below.

We no will proof the general results for estimators with and without transformation

Proof of Theorem 53 The error can be expanded as follows

$$\begin{aligned} U^{\tilde{g}^{N,M}}_{t+\tau,T} &= \frac{\sqrt{M}}{N}\left(N\left(\frac{1}{M}\sum_{i=1}^{M}(g(\tilde{X}^{N,i}_{t+\tau,T}(\tilde{X}^N_{t+\tau})) - g(X^i_{t+\tau,T}(\tilde{X}^N_{t+\tau})))\right)\right) \\ &+\frac{1}{\sqrt{M}}\sum_{i=1}^{M}(g(X^i_{t+\tau,T}(\tilde{X}^N_{t+\tau})) - \mathbf{E}[g(X_{t+\tau,T}(\tilde{X}^N_{t+\tau}))|\mathcal{F}_{t+\tau}]) \\ &+\sqrt{M}\mathbf{E}[(g(X_{t+\tau,T}(\tilde{X}^N_{t+\tau})) - g(X_{t+\tau,T}(X_{t+\tau})))|\mathcal{F}_{t+\tau}] \end{aligned}$$

Given the proof of theorem 48 and the fact that $\mathbf{P} - lim X_{t+\tau}^{N} = X_{t+\tau}$ it only remains to analyze the limit of the last term. By the mean value theorem we can write for some $\lambda_1 \in]0,1[$.

$$\sqrt{M}\mathbf{E}[U_{t+\tau}^{g(X(\tilde{X}^N))}|\mathcal{F}_{t+\tau}] = \sqrt{M}\sum_{j=1}^{d}\mathbf{E}[\partial_j g(\tilde{X}_{t+\tau,T}(\tilde{X}_{t+\tau}) + \lambda_1 e_j \tilde{U}_{j,t+\tau,T}^{X(\tilde{X}^N)})\tilde{U}_{j,t+\tau,T}^{X(\tilde{X}^N)}|\mathcal{F}_{t+\tau}]$$

where $U_{t+\tau,T}^{g(X(\tilde{X}^N))} = g(\tilde{X}_{t+\tau,T}(\tilde{X}_{t+\tau}^N)) - g(X_{t+\tau,T}(X_{t+\tau}))$ and $\tilde{U}_{j,t+\tau,T}^{X(\tilde{X}^N)} = \tilde{X}_{j,t+\tau,T}(\tilde{X}_{t+\tau}^N) - X_{j,t+\tau,T}(X_{t+\tau})$. But

$$\begin{aligned}\tilde{U}_{j,t+\tau,T}^{X(\tilde{X}^N)} &= \tilde{X}_{j,t}^{N} - \tilde{X}_{j,t} \\ &+ \sum_{l=1}^{d}\int_{t+\tau}^{t+\tau}[\partial_l A_j(\tilde{X}_{t,s}(\tilde{X}_{t+\tau}) + \lambda_2 e_l \tilde{U}_{t+\tau,T}^{X(\tilde{X}^N)})ds \\ &+ \sum_{k=1}^{d}\partial_l B_{j,k}(\tilde{X}_{t,s}(\tilde{X}_{t+\tau}) + \lambda_2 e_l \tilde{U}_{t+\tau,T}^{X(\tilde{X}^N)})dW_s^k]\tilde{U}_{l,t+\tau,s}^{X(\tilde{X}^N)}\end{aligned}$$

and consequently

$$\sqrt{M}\mathbf{E}[U_{t+\tau,T}^{g(X(\tilde{X}^N))}|\mathcal{F}_{t+\tau}] = \frac{\sqrt{M}}{\sqrt{N}}\mathbf{E}[\partial g(X_{t+\tau,T})\tilde{\Omega}_{t+\tau,T}|\mathcal{F}_{t+\tau}]U_{t,t+\tau}^{\tilde{X}^N} + o_P(1)$$

as $N \to \infty$. It follows that if $\frac{M}{N} \to \varepsilon^2 < \infty$ as $M, N \to \infty$ we have that

$$\sqrt{M}\mathbf{E}[U_{t+\tau,T}^{g(X(\tilde{X}^N))}|\mathcal{F}_{t+\tau}] \to \varepsilon\mathbf{E}[\partial g(X_{t+\tau,T})\tilde{\Omega}_{t+\tau,T}|\mathcal{F}_{t+\tau}]\tilde{U}_{t,t+\tau}^{X}$$

where $\tilde{\Omega}_{t+\tau,T}$ and $\tilde{U}^X$ are given as in theorem 36. Consequently, since the first term is $O(1/\sqrt{N})$ when $\frac{M}{N}$ is $O(1)$ we have that

$$U_{t+\tau,T}^{\tilde{g}^{N,M}} \Rightarrow M_{t+\tau,T}(X_{t+\tau}) + \varepsilon\mathbf{E}[\partial g(X_{t+\tau,T})\tilde{\Omega}_{t+\tau,T}|\mathcal{F}_{t+\tau}]\tilde{U}_{t+\tau}^{X}$$

Finally, the equivalence with the error corrected estimator follows since as we have seen the first term in the error expansion that converges to the second order bias is asymptotically negligible. ■

Proof of Theorem 54 We can expand the error $U^{\hat{g}^{N,M}}_{t+\tau,T}$ as in the proof of theorem 53 (replace ˜ with ˆ) and obtain for the last term of this expansion that

$$\sqrt{M}\mathbf{E}[U^{\hat{g}(X(\hat{X}^N))}_{t+\tau,T}|\mathcal{F}_t] = \frac{\sqrt{M}}{N}\mathbf{E}[\partial g(X_{t+\tau,T})\hat{\Omega}_{t+\tau,T}|\mathcal{F}_{t+\tau}]U^{\hat{X}^N}_{t,t+\tau} + o_P(1)$$

as $N \to \infty$. It follows that if $\frac{\sqrt{M}}{N} \to \epsilon < \infty$ we have that

$$\sqrt{M}\mathbf{E}[U^{\hat{g}(X(\hat{X}^N))}_{t+\tau,T}|\mathcal{F}_t] \Rightarrow \epsilon\mathbf{E}[\partial g(X_{t+\tau,T})\hat{\Omega}_{t+\tau,T}|\mathcal{F}_{t+\tau}]\hat{U}^X_{t,t+\tau}$$

It follows in this case that the first term will not vanish if $\frac{\sqrt{M}}{N} \to \epsilon < \infty$. Therefore we have in this case

$$U^{\hat{g}^{N,M}}_{t+\tau,T} \Rightarrow \epsilon\hat{K}_{t+\tau,T}(X_{t+\tau}) + M_{t+\tau,T}(X_{t+\tau}) + \epsilon\mathbf{E}[\partial g(X_{t+\tau,T})\hat{\Omega}_{t+\tau,T}|\mathcal{F}_{t+\tau}]\hat{U}^X_{t,t+\tau}$$

The equivalent second order bias corrected estimator with transformation is then

$$U^{\hat{g_c}^{N,M}}_{t+\tau,T} \Rightarrow M_{t+\tau,T}(X_{t+\tau}) + \epsilon\mathbf{E}[\partial g(X_{t+\tau,T})\hat{\Omega}_{t+\tau,T}|\mathcal{F}_{t+\tau}]\hat{U}^X_{t,t+\tau}$$

since as we have seen in this the first term does asymptotically vanish.

■

Proof of Theorem 56 and Theorem57 As in the proof with the known initial value we can use the arguments of the general proof once we have realized that $\frac{\frac{1}{sqrtM}\sum_{i=1}^d H^{N,i}_{j,t+\tau,T}}{\frac{1}{M}\sum_{i=1}^M H^{N,i}_{1,t+\tau,T}} = \frac{\frac{1}{sqrtM}\sum_{i=1}^d H^{N,i}_{j,t+\tau,T}}{\mathbf{E}[\xi^\rho_{t+\tau,T}|\mathcal{F}_{t+\tau}]} + o_P(1)$ for $j = 2,3$ as $M, N \to \infty$. The result then follows from the fact that $H_{j,t+\tau,T}$ is a function of the terminal point of an SDE and the fact that the derivative of this functional with respect to the initial position is as described in definition 55. ■

Chapter 4

Conclusion

The results of the first article can be seen as a necessary step for the construction of general dynamic models with anticipative information. The expressions for optimal portfolio and consumption policies enable us to investigate exactly how pieces of anticipative information affect the optimal behavior of an investor. They can therefore be used to analyze whether or not conclusions about the efficiency of trading mechanisms in the presence of insiders are robust to the probabilistic specification of the model. Furthermore, since the whole setup is non-Markovian, that is prices are not state variables, our results seem more appropriate to address questions about the informational efficiency of prices in financial markets. Such issues are considered in Rindisbacher (1998) where we use the analysis at the individual level of this paper to construct equilibrium models.

The result about the existence of arbitrage opportunities for insiders illustrates that without restrictions on the investment horizon careful modeling of such investors is required. It shows that viability is an important issue for any model of inter-temporal risk sharing.

The techniques introduced in this paper have interesting applications in other fields. In many problems of dynamic risk management it is required to quantify the effects of a worst case scenario on the value of contingent claims and optimal wealth. In this case we can interpret the vector of signals as realizations of such a scenario and the value of insider information as a measure of the value at risk. Such a measure will incorporate the investor's attitude towards risk. This issue will be of interest for future research where we plan to consider enlargements of filtrations not only by random variables but by continuous stochastic processes. Such a generalization will enable us to study exactly the structure of a given flow of information.

In the second paper we have developed a comprehensive approach for the calculation of the optimal portfolio in the asset allocation problem. One major benefit of our method which relies on Monte-Carlo simulation is its flexibility. Indeed the approach can be easily adapted to encompass (i) any finite number of state variables, (ii) any process for the state variables which satisfies the conditions described and (iii) any number of risky assets. It is also valid for any preference relation in the von Neumann-Morgenstern class. This flexibility provides a distinct advantage over alternative approaches to the problem.

The paper also derives a number of economic results which can be used as guidelines for sound asset allocation rules. The lessons drawn from our simulations can be summarized in the following observations:

1. Hedging components cannot be ignored for asset allocation purposes. Even for short investment horizons they imply an adjustment to mean-variance demands which may represent up to 20% of the stock demand. For long investment horizons hedging behavior has a major impact: the adjustment to mean-variance demands can represent up to 80% of the stock demand.

2. Hedging corrections are fairly stable over time: market timing experiments show that the volatility of the hedging components is low relative to the

variation in the mean-variance component.

3. The most important factors in optimal allocation shares are the risk aversion of the investor and the investment horizon. Of particular interest is the behavior of the optimal stock demand relative to the investment horizon, namely the fact that long (short) investment horizons mandate an increase (decrease) in stock holdings relative to myopic behavior. Although this effect was only recorded in the context of our basic bivariate model, it confirms the interest of tailoring investment products and strategies to different categories of clienteles.

4. Allocation shares are also remarkably stable relative to the other parameters of the model. Variations of the order of 2 standard deviations around estimated parameter values have little impact on the magnitude and the behavior of investment shares.

5. In multi-asset models return correlations and correlations between returns and state variables emerge as important factors composition of the optimal portfolio. Even for short horizons asset demands can increase by a factor of 5 when assets returns are highly correlated.

Naturally, the performance of any asset allocation rule will also depend on the soundness of the underlying model of financial markets. Clearly we do not suggest that the models investigated in this paper are adequate in that respect. However, the approach that we have proposed is universal in the sense that it can be used to address the asset allocation problem under complete markets for any realistic specification of the uncertainty structure no matter how complex.

The third paper introduces an asymptotic error analysis for the solution method of the optimal portfolio choice problem presented in the second paper. We present:

- Asymptotic error distributions of the approximation errors for Euler schemes for stochastic differential equations with and without Doss transformation.
- Probabilistic expression of expected approximation errors for both cases.
- Asymptotic distributions of estimators of expected values of functionals of variables satisfying stochastic differential equations.
- Second order bias corrected approximation schemes.
- Asymptotic error distributions for error corrected approximation schemes.
- Asymptotic error distributions of approximation schemes of conditional expectations of functionals of stochastic differential equations with unknown initial value.

Since we are able to characterize the second order asymptotic bias due to the discretization of the stochastic differential equation we are able to develop bias corrected schemes that are asymptotically equivalent to standard Monte Carlo procedures that have been applied in computational finance for a long time. We also show that Doss's transformation which eliminates stochastic integrals will increase the speed of convergence for weak solution schemes of stochastic differential equations but not if these solution schemes have to be combined with a Monte Carlo procedure to estimate a conditional expectation.

Since we first embed the problem in a more general one, our results have important application other domains than portfolio allocation like option pricing of exotic options by Monte Carlo, numerical approximation of stochastic volatility coefficients in general equilibrium models, Error analysis for GARCH models as diffusion approximation.

The three articles together can be seen as on step forward to provide fund managers with tools that allow them to develop decision support systems for optimal dynamic portfolio allocation similar to the ones they already use to trade in a derivatives.

Bibliography

[1] J. Amendinger, P. Imkeller, and M. Schweizer, *Additional logarithmic utility of an insider*, Tech. report, Humboldt Universitaet zu Berlin, 1998.

[2] K. Back, *Asymmetric information and options*, Review of Financial Studies **6** (1993), 435–472.

[3] K. Back and S. Pliska, *On the fundamental theorem of asset pricing with an infinite state space*, Journal of Mathematical Economics **20** (1991), 1–18.

[4] A. Benoussan and R.J. Elliott, *Attainable claims in markov markets*, Math. Finance **5** (1995), 121–131.

[5] M. Broadie Boyle, P. P. and P. Glasserman, *Monte carlo methods for security pricing*, Journal of Economic Dynamics and Control **21** (1997), 1267–1321.

[6] M. Brennan, E. Schwarz, and R. Lagnado, *Strategic asset allocation*, Journal of Economic Dynamics and Control **21** (1997), 1377–1403.

[7] L. Broze, O. Scaillet, and J-M. Zakoian, *Testing for continuous-time models of the short-term interest rate*, Journal of Empirical Finance **3** (1995), 199–223.

[8] J. Campbell and L. Viceira, *Consumption and portfolio decisions when expected returns are time varying*, Quarterly Journal of Economics **114** (1999), 433–495.

[9] K. Chan, G. Karolyi, F. Longstaff, and A. Sanders, *An empirical comparison of alternative models of the short-term interest rate*, Journal of Finance **47** (1992), 1209–1227.

[10] D.B Cowell, R.J. Elliott, and P.E. Kopp, *Martingale representation and hedging policies*, Stoch. Process. Appl **38** (1991), 335–345.

[11] J. C. Cox and C. Huang, *Optimal consumption and portfolio policies when asset prices follow a diffusion process*, Journal of Economic Theory **49** (1989), 33–83.

[12] R.C. Dalang, A. Morton, and W. Willinger, *Equivalent martingale measures and no-arbitrage in stochastics*, Stoch. and Stoch. Rep. **29** (1990), 185–201.

[13] F. Delbaen, P. Monat, W. Schachermayer, M. Schweizer, and C. Stricker, *Weighted norm inequalities closedness of a space of stochastic integrals*, Finance and Stochastics **1** (1997), 25–42.

[14] F. Delbaen and W. Schachermayer, *A general version of the fundamental theorem of asset pricing*, Math. Annalen **30** (1994), 463–520.

[15] ———, *The existence of absolutely continuous local martingale measures*, Annals of Applied Probability **5** (1995), 926–945.

[16] ———, *The fundamental theorem of asset pricing for unbounded stochastic processes*, Tech. report, ETH-Zurich and Universitaet Wien, 1998.

[17] J. Detemple, R. Garcia, and M. Rindisbacher, *A numerical solution for optimal portfolio policies*, Tech. report, CIRANO, 2000.

[18] H. Doss, *Liens entre équations différentielles stochastiques et ordinaires*, Ann. Inst. H. Poincaré **13** (1977), 99–128.

[19] D. Duffie and C. Huang, *Multiperiod security markets with differential information*, Journal of Mathematical Economics **15** (1986), 283–303.

[20] P. Dybvig and C.F. Huang, *Nonnegative wealth, absence of arbitrage, and feasible consumption plans*, Review of Financial Studies **1** (1988), 377–401.

[21] R.J. Elliott, H. Geman, and B.M. Korkie, *Portfolio optimization and contingent claim pricing with differential information*, Stochastics and Stochastic Reports **60** (1997), 185–203.

[22] H. Foellmer and P. Imkeller, *Anticipation cancelled by a girsanov transformation: a paradox on wiener space*, Ann. Inst. Henri Poincaré **29** (1993), 569–586.

[23] E. Fournié, J.M. Lasry, J. Lebuchoux, P.-L. Lions, and N. Touzi, *Some applications of malliavin calculus to monte carlo methods in finance*, Finance and Stochastics **3** (1999), 391–412.

[24] S.J. Grossman, *The informational role of prices*, MIT Press, 1990.

[25] M.J. Harrison and D.M. Kreps, *Martingales and arbitrage in multiperiod securities markets*, Journal of Economic Theory **20** (1979), 381–408.

[26] J. Hirshleifer, *The private and social value of information and the reward to inventive activity*, American Economic Review **61** (1971), 561–574.

[27] P. Imkeller, *Enlargement of wiener filtration by an absolutely continuous random variable via malliavin's calculus*, Probab. Theory and Relat. Fields **106** (1996), 105–135.

[28] J. Jacod, *Calcul stochastique et problème de martingales*, Springer Verlag, 1979.

[29] ———, *Grossissement initial, hypothèse h' et theorème de girsanov*, Lecture Notes in Mathematics, vol. 1118, ch. 3, pp. 15–35, Springer Verlag, 1980.

[30] T. Jeulin and M. Yor., *Inégalité de hardy, semimartingales, et faux-amis*, Lecture Notes in Mathematics, vol. 721, ch. 10, pp. 332–359, Springer Verlag, 1977.

[31] E. Jouini and H. Kallal, *Arbitrage in securities markets with short-sales constraints*, Math. Finance **3** (1995), 178–197.

[32] ———, *Martingales and arbitrage in security markets with transaction costs*, Journal of Economic Theory **3** (1995), 197–232.

[33] I. Karatzas, S. Lehoczky, S. Sethi, and S. Shreve, *Explicit solution of a general consumption/investment problem*, Mathematics of Operation Research **11** (1986), 261–294.

[34] I. Karatzas, S. Lehoczky, and S. Shreve, *Optimal portfolio and consumption decisions for a small investor*, SIAM Journal of Control and Optimization **25** (1987), 1557–1586.

[35] I. Karatzas and D. Ocone, *A generalized clark representation formula, with application to optimal portfolios*, Stochastics **34** (1991), 187–220.

[36] I. Karatzas and I. Pikovsky, *Anticipative porfolio optimization*, Advances in Applied Probability **28** (1996), 1095–1122.

[37] ———, *Market equilibrium with differential information*, Tech. report, Columbia University, 1996.

[38] I. Karatzas and S. Shreve., *Brownian motion and stochastic calculus*, Springer New York, 1987.

[39] P.E. Kloeden and E. Platen, *Numerical solution of stochastic differential equations*, Springer-Verlag, Berlin ; New York, 1992.

[40] D.M. Kreps, *Arbitrage and equilibrium in economies with infinitely many commodities*, Journal of Mathematical Economics **8** (1981), 15–35.

[41] H. Kunita, *Stochastic flows and stochastic differential equations*, Cambridge University Press, Cambridge ; New York, 1990.

[42] T.G. Kurtz and P. Protter, *Wong-zakai corrections, random evolutions and numerical schemes for sde's*, Stochastic Analysis, vol. 10, ch. 7, pp. 331–346, Academic Press, 1991.

[43] A.S. Kyle, *Continuous auctions and insider trading*, Econometrica **53** (1985), 1335–1355.

[44] S. Levental and A.V. Skorohod, *A necessary and sufficient condition for absence of arbitrage with tame portfolios*, Annals of Applied Probability **5** (1995), 906–925.

[45] J. Liu, *Portfolio selection in stochastic environments*, Stanford University, 1999.

[46] A. W. Lo, *Maximum likelihood estimation of generalized ito processes with discretely-sampled data*, Econometric Theory **4** (1988), 231–247.

[47] J. R. Magnus and H. Neudecker, *Matrix differential calculus with applications in statistics and econometrics*, Wiley, New York, 1988.

[48] R.C. Merton, *Lifetime portfolio selection under uncertainty: the continuous-time case*, Review of Economics and Statistics **51** (1969), 247–257.

[49] ———, *Optimum consumption and porftolio rules in a continuous-time model*, Journal of Economic Theory **3** (1971), 373–413.

[50] G. N. Mil'shtein, *Numerical integration of stochastic differential equations*, Kluwer Academic Publisher, Dordrecht ; Boston, 1995.

[51] D. B. Nelson and D. P. Foster, *Asymptotic filtering theory for univariate arch models*, Econometrica **62** (1994), 1–41.

[52] D. Nualart, *Malliavin calculus and related topics*, Springer Verlag, 1995.

[53] H. Pham and N. Touzi, *The fundamental theorem of asset pricing with cone constraints*, Tech. report, CREST Paris, 1996.

[54] D. Revuz and M. Yor., *Continuous martingales and brownian motion*, Springer New York, 1990.

[55] S. F. Richard, *'optimal consumption, portfolio and life insurance rules for an uncertain lived individual in a continuous-time model*, Journal of Financial Economics **2** (1975), 187–203.

[56] M. Rindisbacher, *Insider information and financial market equilirium*, Tech. report, Université de Montréal and CIRANO, 1998.

[57] P.A. Samuelson, *Lifetime portfolio selection by dynamic stochastic programming*, Review of Economics and Statistics **LI** (1969), 239–246.

[58] W. Schachermayer, *Martingale measures for discrete time processes with infinite horizon*, Math. Finance **4** (1994), 25–55.

[59] S. Song, *Grossissements de filtrations et problèmes connexes*, Thèse de doctorat de l'Université de Paris IV, 1987.

[60] C. Stricker and M. Yor, *Calcul stochastique dépendant d'un paramètre*, Z. Wahrscheinlichkeitstheorie verw. Gebiete **45** (1978), 109–133.

[61] D. Talay and E. Pardoux, *Discretization and simulation of stochastic differential equations*, Acta Applicandae Mathematicae **2** (1985), 23–47.

[62] D. Talay and L. Tubaro, *Expansion of the global error for numerical schemes solving stochastic differential equations*, Stochastic Analysis and its Applications **8** (1990), 94–120.

[63] J. Wachter, *Optimal portfolio allocation for an investor with utility over consumption*, Tech. report, Harvard University, 1999.

[64] K.X.D. Wang, *Asset pricing in infinite horizon markets with constraints on porfolio values*, Tech. report, University of Minnesota, 1998.

[65] G.A. Willard and P.H. Dybvig, *Empty promises and arbitrage*, Tech. report, Washington University in Saint Louis, 1996.