Modélisation stochastique des marchés financiers et optimisation de portefeuille

Maxime Bonelli

UNIVERSITÉ DE NICE-SOPHIA ANTIPOLIS - UFR SCIENCES
École Doctorale en Sciences Fondamentales et Appliquées

THÈSE

pour obtenir le titre de

DOCTEUR EN SCIENCES
DE L'UNIVERSITÉ DE NICE-SOPHIA ANTIPOLIS

DISCIPLINE : **MATHÉMATIQUES**

présentée et soutenue par

Maxime BONELLI

MODÉLISATION STOCHASTIQUE DES MARCHÉS FINANCIERS ET OPTIMISATION DE PORTEFEUILLE

Stochastic modeling of financial markets and portfolio optimization

Thèse dirigée par : Mireille BOSSY
soutenue le 8 Septembre 2016

Jury:

Mme Mireille Bossy, Directrice de Recherche INRIA	Directrice
M. Romuald Elie, Professeur à l'Université Paris-Est Marne-la-Vallée	Rapporteur
Mme Nicole El Karoui, Professeur à l'École Polytechnique	Présidente
M. Eric Ghysels, Professeur à l'Université de Caroline du Nord à Chapel Hill	Examinateur
M. Georges Hübner, Professeur à l'Université de Liège HEC-ULg	Rapporteur
M. Lionel Martellini, Professeur à l'EDHEC Business School	Examinateur

INRIA Sophia Antipolis Méditerranée, équipe TOSCA, 06902 SOPHIA ANTIPOLIS

Remerciements

Cette thèse CIFRE a été financée par la société Koris International. En premier lieu, je remercie donc vivement ses dirigeants, Jean-René Giraud et Philippe Malaise, respectivement directeur général et président, pour leur confiance.

Je souhaiterais ensuite remercier ma directrice de thèse Mireille Bossy pour avoir accepté d'encadrer ce projet de recherche. Son accompagnement scientifique et ses nombreux conseils m'ont énormément apporté durant ces trois années. Je lui en suis très reconnaissant.

Mes remerciements sincères vont également à Romuald Elie et Georges Hübner qui ont accepté de rapporter cette thèse en m'apportant de nombreuses remarques et suggestions extrêmement pertinentes. Je remercie respectueusement Nicole El Karoui, Eric Ghysels et Lionel Martellini qui m'ont fait l'honneur d'accepter de faire partie de mon jury de thèse.

Merci également à l'ensemble des membres de l'équipe TOSCA, ainsi qu'à tous les employés de Koris International, avec qui j'ai beaucoup échangé durant ces trois années, tant sur le plan académique que sur celui de l'industrie de la finance. Si cette expérience a été aussi stimulante et captivante c'est aussi grâce à eux.

Enfin, avec de profonds sentiments, je remercie Meryll, mon frère Raphaël et mes parents Sara et Luc, qui, avec foi et conviction, n'ont jamais douté de ma réussite. À cette étape importante de mon parcours, mes pensées vont vers eux.

Résumé

Cette thèse présente trois contributions indépendantes appartenant au domaine de la modélisation stochastique et des mathématiques financières, avec des applications dans le contexte de la gestion de portefeuille.

La première partie de la thèse se concentre sur la modélisation mathématique et l'estimation empirique de la moyenne conditionnelle des rendements du marché actions, aussi appelée rendement espéré du marché, ce qui constitue un problème de premier ordre en finance quantitative. En effet, selon de nombreuses recherches, les rendements espérés, qui sont inobservables en pratique, sont persistants avec un phénomène de retour à la moyenne, et sont donc souvent modélisés à l'aide d'un processus autorégressif AR(1). Cependant, des études empiriques montrent que lors de mauvaises périodes économiques (récessions) la prédictibilité des rendements est plus élevée, i.e., la variance de la moyenne conditionnelle augmente. Étant donné que le modèle AR(1) exclut par construction cette propriété, nous proposons d'utiliser un modèle CIR discrétisé pour les rendements espérés. Ce dernier conserve le coefficient de dérive de l'AR(1) mais introduit une variance variable au cours du temps qui augmente avec le niveau du processus. Les implications de cette spécification sont étudiées dans le cadre flexible d'un modèle espace-état bayésien.

La deuxième partie de la thèse est dédiée à l'étude et la modélisation de la dynamique jointe de la volatilité des actions et des volumes de transaction. La relation empirique entre ces deux quantités a été justifiée par des modèles théoriques, tels que l'hypothèse de mélange de distribution (MDH). Cette spécification souligne que l'activité transactionnelle et la volatilité des actions sont étroitement liées au taux latent d'arrivée d'information, impliquant une dynamique jointe. Cependant, ce cadre de modélisation ne capture notamment pas la persistance évidente de la variance des actions, à la différence des spécifications GARCH. Nous proposons un modèle de volatilité à deux facteurs combinant les deux approches. Contrairement aux modèles MDH standards, la variance conditionnelle est gouvernée par le processus stochastique d'arrivée d'information, ainsi que par une composante GARCH persistante et avec un retour à la moyenne, afin de dissocier les variations de volatilité court terme et long terme. Le modèle révèle plusieurs régularités importantes sur la relation volume-volatilité en ligne avec les observations empiriques.

La troisième partie de la thèse s'intéresse à l'analyse des stratégies d'investissement optimales sous la contrainte que la valeur du portefeuille ne descende pas en dessous d'une fraction fixée de son maximum courant. Le problème étudié est celui de la maximisation d'utilité à horizon fini pour différentes fonctions d'utilité (concave et asymétrique). Nous calculons les stratégies optimales en résolvant l'équation de Hamilton-Jacobi-Bellman, qui caractérise le principe de programmation dynamique relié au problème de contrôle stochastique. Le problème est résolu numériquement avec une condition de bord qui caractérise la contrainte de perte. En se basant sur un large panel d'expérimentations numériques, nous obtenons numériquement les allocations optimales et nous analysons leurs divergences, en faisant varier les paramètres du modèle de marché et les profils d'utilité des investisseurs.

Mots clés : marché financier, modèle espace-état, filtre de Kalman, analyse bayésienne, volatilité stochastique, modèle GARCH, optimisation de portefeuille, contrôle stochastique, finance comportementale.

Stochastic modeling of financial markets and portfolio optimization

Abstract

This PhD thesis presents three independent contributions related to stochastic modeling and financial mathematics, with applications within the context of portfolio management.

The first part of the thesis is concentrated on the mathematical modeling and the empirical estimation of the conditional mean of stock market returns, also called expected market return, which is a first order issue in quantitative finance. Indeed, a large body of research supports the view that the unobservable expected return process is persistent and mean-reverting, and therefore often modeled as an Auto-Regressive process of order one: AR(1). However, empirical studies have found that during bad times (economic recessions) return predictability is higher, i.e., the variance of the conditional mean increases. Given that the AR(1) model excludes by construction this property, we propose to use instead a discretized CIR model for expected returns. The latter preserves the same drift as the AR(1), but induces a time-varying variance which increases with the level of the process. The implications of this specification are studied within a flexible Bayesian state-space model.

The second part of this dissertation is dedicated to the study and modeling of the joint dynamics of stocks volatility and trading volume. The empirical relationship between these two quantities has been justified by theoretical models, such as the Mixture of Distribution Hypothesis (MDH). This specification predicts that both trading activity and equity volatility are closely linked to the latent information arrival rate, implying a joint dynamics. However, this framework notably fails to capture the obvious persistence in stock variance, unlike GARCH specifications. We propose a two-factor model of volatility combining both approaches. Unlike typical MDH models, the conditional variance is governed by the stochastic information arrival as well as a persistent mean-reverting GARCH component, in order to disentangle short-run from long-run volatility variations. The model reveals several important regularities on the volume-volatility relationship, in line with empirical observations.

The third part of the thesis is concerned with the analysis of optimal investment strategies under the drawdown constraint that the wealth process never falls below a fixed fraction of its running maximum. The finite horizon expectation maximization problem is studied for different types of utility functions (concave, i.e., power utility, and asymmetric, i.e., S-shape). We compute the optimal investments strategies, by solving the Hamilton–Jacobi–Bellman equation, that characterizes the dynamic programming principle related to the stochastic control problem. The problem is numerically solved with a boundary condition that characterizes the drawdown constraint. Based on a large panel of numerical experiments, we compute numerically optimal allocation programs and analyze their divergences, making the market model parameters and investor utility profiles vary.

Keywords : stock market, state-space model, Kalman filter, Bayesian analysis, stochastic volatility, GARCH model, portfolio optimization, stochastic control, behavioral finance.

Contents

Introduction générale **1**

1 An Alternative Model of Expected Returns and its Implications for Return Predictability **7**

1.1 Introduction . . . 8
1.2 A predictive system with heteroscedastic expected returns . . . 10
1.2.1 Modeling expected returns . . . 10
1.2.2 The CIR predictive system . . . 12
1.2.3 Structural differences between the CIR and AR systems . . . 14
1.2.4 Economic Implications . . . 20
1.3 Empirical analysis . . . 23
1.3.1 What are plausible values for R^2, β and E_r? . . . 24
1.3.2 Out-of-sample return prediction using point estimate parameters . . . 26
1.3.3 Out-of-sample return prediction using Bayesian parameters estimates 27
Conclusion . . . 32
Tables and Figures . . . 34
Appendices . . . 45
1.A Discretization of CIR process . . . 45
1.B Autocovariance of returns . . . 46
1.C The Kalman filter . . . 49
1.C.1 The algorithm . . . 50
1.C.2 Steady state . . . 52
1.D Bayesian procedure . . . 53
1.D.1 Drawing $\mu_{\mathbf{t}}$ . . . 54
1.D.2 Prior distributions . . . 54
1.D.3 Posterior distributions . . . 54

2 Stock Market Volatility Dynamics: A Volume Filtered-GARCH Model **57**

2.1 Introduction . . . 58
2.2 A two-component volatility model . . . 61
2.2.1 The mixing variable as stochastic volume component . . . 63
2.2.2 An additional persistent factor in volatility . . . 66
2.3 Properties of the return distribution . . . 67
2.3.1 Conditional expectation of return . . . 68

2.3.2 Higher moments . . . 68
2.3.3 Covariance between return and volume . . . 70
2.4 Empirical analysis . . . 71
2.4.1 Data . . . 72
2.4.2 Maximum Likelihood Estimation . . . 72
2.4.3 Model selection procedure . . . 73
2.4.4 Fitted parameters and conditional estimates . . . 75
2.5 Structural breaks and sub-samples . . . 79
2.5.1 Structural breaks in volatility . . . 79
2.5.2 Fitted parameters . . . 80
2.5.3 Volatility drivers . . . 81
2.6 The effect of unexpected volume on volatility . . . 81
Conclusion . . . 83
Tables and Figures . . . 85

3 Portfolio Management with Drawdown Constraint: An Analysis of Optimal Investment 103
3.1 Introduction . . . 104
3.2 The model . . . 107
3.3 Portfolio management with drawdown constraint . . . 108
3.4 The Hamilton-Jacobi-Bellman equation . . . 109
3.5 Heuristic for the optimal control study . . . 111
3.5.1 Pontryagin maximum principle . . . 111
3.5.2 Intuition on the allocation profiles . . . 111
3.6 Numerical resolution . . . 113
3.7 Empirical analysis based on simulated markets . . . 114
3.7.1 Stochastic sampling of the markets and portfolios . . . 114
3.7.2 Empirical results . . . 117
3.7.3 Prospect theory investor: impact of a higher reference point . . . 121
3.7.4 Mis-estimation risk: impact on preference . . . 123
3.7.5 Shorter investment horizon: impact on preference . . . 125
Conclusion . . . 127
Tables and Figures . . . 129
Appendices . . . 166
3.A Discretization method . . . 166
3.B Wealth process simulation . . . 167
3.C Results with $\alpha = 0.5$ . . . 167

Introduction générale

Cette introduction générale a pour but de présenter la thématique des travaux de recherche de la thèse, ainsi que sa structure composée de trois contributions indépendantes.

La complexité et l'incertitude autour de l'interaction des agents et de l'environnement économique ont amené la finance à être étudiée dans un cadre qui requiert des outils analytiques et mathématiques sophistiqués[1]. Il existe, plus ou moins distinctement, deux branches de recherche en mathématiques financières avec des objectifs et des outils de modélisation différents[2]. La première, initiée par les travaux de Bachelier (1900), est principalement concentrée sur la modélisation de processus en temps continu. À l'aide d'outils tels que le calcul d'Itô et les équations aux dérivées partielles, cette branche conduit à des résultats concernant notamment l'évaluation des produits dérivés et même des résultats mathématiques s'inscrivant dans un cadre plus général tel que la théorie des probabilités et l'optimisation. La deuxième branche est majoritairement focalisée sur l'estimation des risques et la gestion de portefeuille. Cette approche fait appel à des notions appartenant aux statistiques et à l'économétrie, et utilise le plus souvent une modélisation discrète basée sur les séries temporelles. Cette thèse se situe à la frontière des deux domaines décrits ci-dessus. En effet, bien que ces deux branches soient effectivement distinctes, notamment en termes de publications et de littérature, il n'en reste pas moins que les processus stochastiques et les méthodes numériques constituent les briques de modélisation communes aux deux approches. En ce sens, cette thèse traite trois sujets de recherche qui constituent des travaux motivés par des applications pratiques futures en finance. Les contributions s'inscrivent donc dans une littérature tournée vers la finance quantitative directement applicable dans le cadre de la gestion d'investissements, plutôt que vers des résultats mathématiques fondamentaux dans un cadre plus général.

La motivation initiale de cette thèse était celle d'explorer des notions alternatives au cadre standard de modélisation et d'estimation utilisé en finance. Les travaux résultants sont ainsi caractérisés par l'étude de systèmes stochastiques sous des hypothèses nouvelles, avec pour but de répondre à plusieurs problématiques. Quelles sont les implications que peuvent avoir des hypothèses alternatives sur les processus générateurs habituellement utilisés dans les modèles financiers ? Comment réconcilier deux familles de modèles qui décrivent différemment l'évolution d'un même type de processus ? Les résultats empiriques décrits par des recherches en finance se vérifient-ils dans un cadre d'optimalité au sens mathématique ?

[1] Voir par exemple Merton (1995) sur ce sujet.

[2] Meucci (2011) fournit un aperçu d'un point de vue pratique de cette distinction.

Cette thèse se compose de trois contributions indépendantes, faisant chacune l'objet d'un chapitre dans ce manuscrit, décrits successivement ci-après. Les deux premiers chapitres ont pour but de contribuer à la compréhension de la distribution et des caractéristiques de la dynamique des rendements des actions. On s'intéresse à ce type d'actif, primordial en finance, à travers la modélisation du processus correspondant au premier moment dans le premier chapitre (moyenne conditionnelle ou rendement espéré), puis du processus correspondant au second moment (variance conditionnelle) dans le deuxième chapitre. Le troisième et dernier chapitre est, quant à lui, focalisé sur la thématique de l'optimisation de portefeuille sous contraintes de pertes maximales.

Modélisation et estimation des rendements espérés du marché actions

Le premier chapitre de la thèse est focalisé sur une question de premier ordre en finance quantitative : comment modéliser et estimer la moyenne conditionnelle des rendements des actions (non observable en pratique), aussi appelée rendement espéré des actions et dénotée μ_t ci-après ? Cette question est importante car elle est directement liée à la notion de prédictibilité dans le marché actions : une moyenne conditionnelle constante est synonyme de marche aléatoire, alors qu'une moyenne conditionnelle variable implique un niveau de prédictibilité non nul. La question de la modélisation des rendements espérés a été en premier lieu traitée dans le cadre de la régression prédictive. Ce cas consiste à utiliser plusieurs variables économiques comme prédicteurs linéaires des rendements futurs. Plusieurs modèles de régression ont été proposés mais le fait que la régression admette une corrélation parfaite entre les prédicteurs et les rendements espérés est une limitation importante. Pástor and Stambaugh (2009) ont proposé un cadre différent, appelé système prédictif, correspondant à un modèle espace-état. Dans ce système on admet que les rendements espérés constituent une variable latente, suivant un processus autorégressif d'ordre un AR(1) inobservable dont les innovations sont potentiellement corrélées à celles des variables prédictives observables. Ce système, beaucoup plus flexible que la régression linéaire, permet d'estimer μ_t à l'aide d'un filtre de Kalman en capturant l'idée que les rendements espérés sont persistants avec un retour à la moyenne.

Cependant le modèle AR(1) exclut par construction une variance variable pour le rendement espéré, et donc un degré de prédictibilité variable au cours du temps ne peut pas être modélisé par ce modèle. Étant donné que plusieurs études empiriques supportent l'hypothèse d'un niveau de prédictibilité variable, nous proposons d'utiliser un modèle alternatif dans le cadre du système espace-état pour représenter le processus inobservable des rendements espérés. Ainsi, nous utilisons à la place du modèle AR(1), le modèle CIR discrétisé. Ce dernier conserve le même coefficient de dérive que le processus AR(1), permettant de capturer le caractère persistant avec retour à la moyenne, mais introduit une variance variable au cours du temps qui augmente avec le niveau du processus. En conséquence, selon ce modèle, si les rendements espérés sont plus élevés pendant les périodes de récession (contracyclicité) ils sont également plus volatils et les rendements réalisés plus

prévisibles. En se basant sur cette nouvelle modélisation, nous estimons μ_t à l'aide d'un filtre de Kalman étendu. La contribution du premier chapitre est l'étude de ce nouveau système et de ses implications sur la dynamique des actions.

Le système CIR introduit est examiné en plusieurs étapes. En premier lieu, le nouveau modèle espace-état est étudié de manière analytique, en mettant en avant ses différences structurelles avec le modèle AR(1). Notamment, le système CIR induit une autocovariance variable au cours du temps pour les rendements, proportionnelle au niveau de μ_t. Ainsi, durant les phases de récessions, les espérances de rendements sont plus élevées, les rendements plus fortement autocorrélés et prévisibles. Nous présentons également les autres conséquences et propriétés du modèle impliquées notamment par l'hétéroscédasticité du processus CIR.

Nous nous focalisons ensuite sur une analyse empirique utilisant les données du marché actions américain sur plusieurs décennies avec une fréquence trimestrielle, c'est-à-dire un horizon relativement long avec une basse fréquence. Nous montrons d'abord comment le système CIR, sans aucun prédicteur, est capable de reproduire naturellement le comportement de certains estimateurs de rendements espérés utilisant la régression prédictive. Nous étudions ensuite le système hors-échantillon à l'aide d'une méthode d'estimation bayésienne. Le système CIR génère des prédictions de rendement significativement meilleures par rapport à une moyenne d'échantillon, et améliore également la performance par rapport au modèle AR(1) et à la régression prédictive utilisant le taux de dividende. De plus, notre étude empirique met en avant l'évolution temporelle de certains paramètres majeurs du modèle à travers leur distribution postérieure bayésienne au cours du temps. L'estimation révèle que le paramètre de persistance des rendements espérés a diminué au cours des deux dernières décennies, de même que l'autocorrélation empirique des rendements réalisés, ce qui coïncide avec une diminution du niveau de prédictibilité.

Le papier de recherche correspondant à ce premier chapitre est le suivant :
Bonelli, M., et Mantilla-Garcia, D. (2016). *An Alternative Model of Expected Returns and its Implications for Return Predictability.* Disponible sur SSRN.
Il a été présenté aux conférences internationales suivantes avec comité de sélection avant acceptation :

- Annual Meeting European Financial Management Association (EFMA) 2015 (nominé pour le *conference best paper award*),
- International Conference of the Financial Engineering and Banking Society (FEBS) 2015,
- International Conference of the French Finance Association (AFFI) 2015,
- International Forecasting Financial Markets Conference (FFM) 2015,
- International Conference on Computational and Financial Econometrics (CFE) 2014.

Modélisation de la volatilité des actions et de sa relation avec les volumes de transaction

Le deuxième chapitre se concentre sur l'étude de la dynamique de la volatilité des rendements des actions, et plus précisément sur sa relation avec les volumes de transaction, i.e., le nombre de titres échangés chaque jour. Plusieurs modèles ont été proposés afin de justifier théoriquement le lien empirique observé entre les mouvements des prix et les volumes. En particulier, les modèles de mélange de distribution (MDH) avancent que la volatilité des actions et la quantité de titres échangés sont liées au taux latent d'arrivée d'information, ce qui engendre une dynamique jointe. Cependant ce type de modélisation ne capture pas certaines caractéristiques importantes de la volatilité, en particulier sa persistance. Ainsi, nous proposons un modèle à deux composantes dans le but de dissocier les mouvements court terme de volatilité potentiellement dus à l'arrivée de nouvelles informations des mouvements long terme et persistants de la volatilité.

La modélisation introduite ici est paramétrique et implique que la volatilité des rendements est gouvernée à la fois par le flux stochastique d'arrivée d'information (identifiée par une fonction non linéaire des volumes) et par une composante de type GARCH persistante sans rapport avec les volumes de transaction. Le modèle permet ainsi d'avoir une partie de la persistance de la volatilité explicitement non liée à celle du taux d'arrivée d'information. De cette manière le modèle est capable d'expliquer conjointement la persistance de la volatilité et sa relation avec les volumes. Sur ce dernier point, la composante volume du modèle est décomposée en une partie attendue et une partie inattendue (innovations). Cette séparation permet de montrer que la composante stochastique de la volatilité est étroitement liée aux innovations des volumes de transaction. Notons par ailleurs que la spécification globale introduite implique une distribution conditionnelle normale log-normale asymétrique avec queues épaisses pour les rendements.

La contribution de ce deuxième chapitre est l'analyse du modèle ainsi que de ses conséquences à l'aide d'une estimation empirique basée sur le maximum de vraisemblance. En utilisant les données journalières d'actions américaines sur les vingt dernières années, nous montrons que les volumes attendus et inattendus ont chacun un effet opposé, réduisant et augmentant respectivement la volatilité. Par ailleurs, on obtient que la persistance de la volatilité (mise en avant par les modèles de type GARCH) n'est pas réduite dans notre modèle : le caractère persistant de la volatilité est principalement sans rapport avec la composante volume. Enfin, pour la plupart des titres étudiés, le modèle révèle qu'une grande partie des variations de la volatilité est due à la composante volume et plus spécifiquement aux innovations dans les volumes. La majorité des variations restantes sont expliquées par la composante GARCH persistante et donc sans rapport avec les volumes de transaction.

Le papier de recherche correspondant à ce deuxième chapitre est le suivant :
Bonelli, M. (2016). *Stock Market Volatility Dynamics: A Volume Filtered-GARCH Model.* Disponible sur SSRN.

Il a été présenté aux conférences internationales suivantes avec comité de sélection avant acceptation :

- International Conference of the French Finance Association (AFFI) 2016,
- European Winter Meetings of the Econometric Society 2015.

Analyse de stratégies optimales sous contrainte de pertes cumulées

Le troisième chapitre de la thèse est focalisé sur l'analyse de stratégies d'investissement optimales dans le cadre de l'assurance de portefeuille. Spécifiquement, on considère le problème de maximisation d'utilité à horizon fini sous la contrainte que la richesse ne tombe jamais en dessous d'une certaine fraction de son maximum courant. L'objectif est d'analyser les comportements impliqués par les allocations optimales et surtout d'étudier la préférence pour ce type de contraintes par différents investisseurs.

Les stratégies dites d'assurance de portefeuille peuvent apparaître attractives car elles sont conçues pour protéger les actifs sous gestion et en même temps permettre de participer aux phases de hausses des actifs risqués tels que les actions. Plusieurs auteurs ont étudié la popularité de ces produits et le profil des investisseurs qui pourraient en bénéficier. Cependant étant donné la difficulté de justifier à partir de la théorie d'utilité standard (i.e., fonction d'utilité concave) la popularité de l'assurance de portefeuille observée en pratique, plusieurs études ont proposé de répondre à cette problématique dans le cadre de la finance comportementale. Le résultat est que les investisseurs avec une fonction d'utilité asymétrique entre les gains et les pertes par rapport à un point de référence ("S-shape") pourraient potentiellement préférer les assurances de portefeuille plutôt que d'autres stratégies standards, en terme de maximisation d'utilité. Cependant ces études ne considèrent pas les stratégies dynamiques optimales du point de vue de l'objectif de la maximisation d'utilité. En effet, les contrôles constants utilisés (paramètre clé de la stratégie) peuvent diverger des contrôles optimaux variables au cours du temps. Ainsi, les résultats obtenus à l'aide de ces derniers pourraient être différents de ceux précédemment cités.

Nous analysons les programmes optimaux d'allocation sous la contrainte de perte cumulée, dite contrainte "drawdown". Pour cela, nous nous basons sur un large panel d'expérimentations numériques produites à partir de la formulation du problème de contrôle stochastique correspondant. Le problème considéré est celui de la maximisation de l'utilité espérée à horizon fini pour deux types d'investisseurs, l'un avec fonction d'utilité concave, l'autre avec fonction d'utilité asymétrique. Les stratégies optimales sont dérivées en résolvant l'équation de Hamilton-Jacobi-Bellman caractérisant le principe de programmation dynamique relié à notre problème. L'équation aux dérivées partielles résultante est résolue numériquement sur un espace discrétisé, en faisant varier les paramètres du modèle de marché financier et les profils d'utilité.

A l'aide de simulations stochastiques, nous montrons que, d'un point de vue optimal,

une explication claire sur l'attractivité des stratégies d'assurance de portefeuille ne peut pas être tirée en se basant sur l'investisseur avec utilité asymétrique. D'une part, pour l'investisseur standard (utilité concave), la préférence pour la stratégie optimale avec la contrainte drawdown plutôt que la stratégie optimale sans protection dépend fortement des paramètres de la dynamique de l'actif risqué. D'autre part, nous trouvons que l'investisseur avec utilité asymétrique préfère la stratégie optimale sans contrainte. Ce résultat peut être expliqué par la présence du point de référence dans la fonction d'utilité asymétrique, qui agit comme une protection implicite. Ceci implique que la contrainte n'est pas nécessairement utile pour ce type d'utilité. Les seules configurations où la contrainte de perte se révèle profitable pour cet investisseur correspondent aux situations où nous imposons une erreur d'estimation sur les paramètres du modèle. Dans ces cas spécifiques, la limitation des pertes à travers la contrainte peut jouer le rôle d'assurance contre le risque d'estimation du modèle.

Le papier de recherche correspondant à ce troisième chapitre est le suivant :
Bonelli, M., et Bossy, M. (2016). *Portfolio Management with Drawdown Constraint: An Analysis of Optimal Investment*. En préparation.

Chapter 1

An Alternative Model of Expected Returns and its Implications for Return Predictability

Empirical research on financial markets has found that the conditional mean of the market return process (a.k.a., expected return) is persistent and mean-reverting, and therefore often modeled as an autoregressive process of order one: AR(1). However, several recent studies suggest that during bad times return predictability is higher. Thus, variation in the conditional expected return process should be relatively higher as economic conditions worsen. Hence, we propose instead to model expected returns as a CIR process, which is a parsimonious specification that captures the countercyclical dynamics of stock return predictability. The implications of the model are studied within a flexible Bayesian state-space system. In our empirical tests, the estimates from the proposed model of expected return present lower prediction errors than the historical mean, and than the predictive regression using the dividend yield (which is a well known economic predictor). Furthermore, according to posterior distributions, the persistence parameter of expected returns presents long-term dynamics similar to the autocorrelation of realized returns.

Keywords: return predictability, Kalman filter, Bayesian analysis.

1.1 Introduction

How should we model and estimate the expected excess return of stocks[1]? This question has primarily been addressed within the linear predictive regression framework (early examples are Keim and Stambaugh, 1986; Stambaugh, 1986; Ferson and Harvey, 1991; Pesaran and Timmermann, 1995; Stambaugh, 1999). Several economic predictors have been investigated using the predictive regression, including the dividend yield (Fama and French, 1988; Campbell and Shiller, 1988; Goyal and Welch, 2003; Ang and Bekaert, 2007; Lettau and van Nieuwerburgh, 2008; Cochrane, 2008), the interest rates (e.g. Campbell, 1987), the term and default spreads (Campbell, 1987; Fama and French, 1989), variables including information of corporate payout and financing activity (e.g. Baker and Wurgler, 2000), the consumption-wealth ratio (Lettau and Ludvigson, 2001), and combinations of these (Goyal and Welch, 2008; Rapach, Strauss, and Zhou, 2010).

The idea that high valuation ratios, such as dividend-price and earnings-price ratios should predict high subsequent returns dates back to the tradition of value investors of Graham, Dodd, and Cottle (1934). Since then, the most pervasive predictor found in the literature is the dividend yield, backed also by the famous log linearization of the dividend-price ratio in Campbell and Shiller (1988), which implies a positive linear relationship between expected returns and the ratio. The dividend-price ratio is a relatively stable[2], persistent, approximately mean-reverting quantity that is naturally counter-cyclical, i.e., it is high when prices are historically low and vice-versa. Arguably, the direct relation with the valuation ratio has to some extent led many financial economists to hold the view that expected returns should have those characteristics: stable, persistent, mean-reverting and counter-cyclical.

Advanced regression models have been proposed in the literature to analyze expected returns (see Dangl and Halling, 2012; Johannes, Korteweg, and Polson, 2014; Pettenuzzo and Ravazzolo, 2014), but the regression approach has some limits, including the fact that it assumes a perfect correlation between the predictor(s) and expected returns. Pástor and Stambaugh (2009) introduced a *predictive system* that allows for imperfect correlations between predictors and expected returns, as well as amongst past realized returns and expected returns. Pástor and Stambaugh (2009) modeled expected returns as an autoregressive process (AR) of order one[3], under a set of informative priors that reflects the belief that they should have those characteristics: stability, persistence, mean-reversion and counter-cyclicality.

However other characteristics supported by empirical observations and economic theory are absent in the AR(1)-based system. In particular, this model excludes by construction a time-varying degree of return predictability. Therefore, we propose to use instead a discretization of the Cox, Ingersoll Jr, and Ross (1985) (CIR) process to model the

[1]Unless explicitly stated otherwise, we use the term *expected returns* as shorthand for *expected excess returns of the stock market* over the risk-free rate or *equity risk premium*. Similarly, *realized returns* stands for *realized excess market returns*.

[2]Relatively stable meaning less variable than realized returns.

[3]Binsbergen, Jules, and Koijen (2010), Pástor and Stambaugh (2012), and Carvalho, Lopes, and McCulloch (2015) also consider this specification.

dynamics of expected returns within the predictive system. This model keeps the four AR(1) relevant features mentioned but captures some additional important characteristics of expected returns. Notably, the CIR system yields a countercyclical predictive power, it implies a time-varying conditional autocovariance of returns (higher in absolute value during bad times) and a positive relation between expected returns and return variance.

The CIR dynamics induce a continuously changing conditional variance for expected returns, which increases with the level of expected returns during market downturns. A time-varying variance in expected returns is motivated by the fact that the predictability of returns increases during economic recessions (see Rapach et al., 2010; Henkel, Martin, and Nardari, 2011; Cujean and Hasler, 2014). Return predictability is measured as the fraction of the variance in realized returns explained by variations in expected returns, i.e., the R^2 of the regression of realized returns on expected returns. It follows that the variance of expected returns must increase during economic recessions, to account for the increase in R^2 during these periods[4]. We find that the conditional heteroscedasticity of expected returns in the CIR system (without any predictor) reproduces the countercyclical predictability of the dividend yield in the predictive regression documented in former studies such as Rapach et al. (2010), Mantilla-Garcia and Vaidyanathan (2011), and Henkel et al. (2011). Other models featuring time-varying variance for expected returns are studied in Pástor and Stambaugh (2012); Bollerslev, Xu, and Zhou (2015); Piatti and Trojani (2015), however our approach has the advantage to be a very parsimonious departure from the AR(1) assumption, without any additional parameter.

Furthermore, the CIR system implies a time-varying conditional covariance between expected and realized returns. The strength of this relationship depends as well on the level of the expected return. For countercyclical expected returns, the CIR system implies that the covariance is higher (in absolute value) when the expected return is high, i.e., during bad times.

Similarly, the CIR model introduces a time-varying conditional autocovariance of returns. The conditional autocovariance of returns is proportional in absolute value to the level of expected returns. Thus, during recessions when return expectations are high and highly volatile, returns are more strongly autocorrelated and predictable. In contrast, when the expected return level and their volatility are low, returns are conditionally less autocorrelated, thus locally less persistent, as the mean reversion of expected returns is stronger.

Most standard asset pricing models imply a positive relation between expected returns and return variance. Unlike the AR(1), the CIR model for expected returns incorporates a positive relationship between its level and its conditional variance, which in turn implies a positive relation between the level of expected returns and the expected variance of realized returns. Furthermore, the CIR-based expected return process has a negligible probability of being negative. This additional feature is compatible with economic intuition: risk-

[4]Even if during market downturns the variance of realized returns increases (see Schwert, 1989; Hamilton and Lin, 1996; Ait-Sahalia and Kimmel, 2007), the variance of expected returns must increase to compensate for the variance rise of realized returns combined with a higher R^2.

averse investors would not hold stocks if the equity premium was negative[5]. The positivity of the system is also in line with empirical results by Campbell and Thompson (2008), and Pettenuzzo, Timmermann, and Valkanov (2014) who find that restricting expected return estimates to be positive improves their robustness out-of-sample, within the predictive regression framework[6].

In our empirical analysis we find that the CIR predictive system without any predictors (i.e., using only past realized returns) can produce return forecasts out-of-sample that are significantly better than the historical average and the predictive regression using the dividend yield as the predictor, if economically motivated priors are used for the parameters of the system. The CIR system produces better forecasts under the priors that expected returns are countercyclical and display a relatively low variance (i.e., stable), which is consistent with widely held beliefs about expected returns. On the other hand, in line with Kelly and Pruitt (2013), according to the CIR system the persistence parameter of expected returns is lower, i.e., less close to 1, than suggested by the AR system and previous studies. We find that its value in the CIR system has been declining over the last thirty years in tandem with the sample autocorrelation of realized returns, which contrasts with the high and stable persistence parameter estimate of the AR system.

1.2 A predictive system with heteroscedastic expected returns

In this section we present a short summary of common predictive systems in the literature and then introduce our model along with its theoretical implications.

1.2.1 Modeling expected returns

The discrete dynamics for the realized return r at time $t+1$ can be decomposed as

$$r_{t+1} = \mu_t + u_{t+1}, \tag{1.1}$$

where μ_t is the expected return at time t, and the innovation u_{t+1} is the "unexpected return", which is a random shock with zero mean and variance σ_u^2.

In the classic random walk model, expected returns are assumed to be a constant, and the estimate of μ is usually the sample average of all the history of returns available. The random walk model is the no-predictability benchmark that a predictive system with time-varying expected returns has to outperform in terms of forecasting error.

The most common predictive system in the literature is the predictive regression, in which the expected return process μ is assumed to be a linear function of a set of predictors x,

$$\mu_t = a + b'x_t \tag{1.2}$$

[5]Merton (1980) estimates instantaneous expected return on the market and concludes that: *"in estimating models of the expected market return, the non-negativity restriction of the expected excess return should be explicitly included as part of the specification"* (Merton, 1980, p. 323).

[6]In the CIR system, the positivity condition is not imposed as a constraint but it is a result of the dynamics of the expected returns process.

for constants a and b with suitable dimensions. Notice that the dynamics of μ in this model are completely determined by the dynamics in the predictor variables.

The expected return process μ is generally believed to be a mean-reverting process, varying along its long-term average E_r. For example the predictive system of Pástor and Stambaugh (2009) is a state-space model in which the latent expected return process follows a first-order autoregressive AR(1) equation:

$$\mu_{t+1} = (1-\beta)E_r + \beta\mu_t + w_{t+1}, \tag{1.3}$$

where E_r denotes the unconditional expectation of r, which is equal to the unconditional expectation of μ and is constant over time; β is a constant persistence parameter assumed to be within $(0,1)$ so that μ is stationary, and w_{t+1} is the innovation in the expected return. We refer to the predictive system introduced by Pástor and Stambaugh (2009) as the AR (autoregressive) system.

In addition, Pástor and Stambaugh (2009) also consider a set of stationary (observable) predictors x_t following a first-order vector autoregressive VAR(1) process,

$$x_{t+1} = (I-A)E_x + Ax_t + v_{t+1}, \tag{1.4}$$

where E_x is the unconditional expectation of x, A is a matrix with suitable dimensions containing the autoregressive coefficients and with eigenvalues lying inside the unit circle, and v is gaussian noise. Furthermore, the three innovation processes above are assumed to be correlated white-noise, independent and identically distributed across t as,

$$\begin{bmatrix} u_t \\ v_t \\ w_t \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_u^2 & \sigma_{uv} & \sigma_{uw} \\ \sigma_{vu} & \Sigma_{vv} & \sigma_{vw} \\ \sigma_{wu} & \sigma_{wv} & \sigma_w^2 \end{bmatrix} \right). \tag{1.5}$$

Denote the covariance matrix in (1.5) as Σ. Notice that the interaction between predictors and expected returns happens trough the correlation between their corresponding innovations v and w.

While in the standard predictive regression the correlation between the predictor(s) and expected returns is assumed to be perfect, the predictive system (1.1), (1.3), (1.4), (1.5) allows for "imperfect predictors" presenting a correlation with expected returns lower than 1 in magnitude.

Pástor and Stambaugh (2009) showed that the standard predictive regression is a particular case of the predictive system in which the correlation between the innovation in the predictor x and innovation in μ is assumed to be perfect, i.e., $\rho_{vw} = \pm 1$, and the autoregressive coefficient of μ and x are equal, e.g., $\beta = A$ if we consider one predictor.

Another distinctive characteristic of this predictive system is that expected returns are also linked with past realized returns through the correlation between their innovations, i.e., ρ_{uw}. This correlation has an impact on the relationship between expected returns and past realized returns, and on the relative importance of what Pástor and Stambaugh (2009) called the *level effect* and the *change effect* in the system. The level effect captures

the procyclicality of expected returns, i.e., the extent to which observing relatively higher (lower) realized returns is a signal of higher (lower) expected returns, while the change effect captures the extent to which observing relatively higher (lower) realized returns is a signal of lower (higher) expected returns (countercyclicality). For the change effect to dominate, ρ_{uw} must be sufficiently negative. If the change effect prevails then expected returns are countercyclical. Pástor and Stambaugh (2009) argue that the change effect should dominate the level effect and other studies such as Campbell (1991), Campbell and Ammer (1993) and Binsbergen et al. (2010) point in the same direction.

Equations (1.1), (1.3), (1.4) and (1.5) constitute a state-space model[7] in which we have $\mathbb{E}(\mu_t|D_t) = \mathbb{E}(r_{t+1}|D_t)$, where D_t denotes the information set available at time t of observable quantities r and x. Hence, a linear Kalman filter can be used to estimate the unobservable expected return process[8] μ.

In the Bayesian empirical analysis of their predictive system, Pástor and Stambaugh (2009) used prior distributions of the input parameters of the system reflecting "*the prior belief that the conditional expected return μ_t is stable and persistent*" (Pástor and Stambaugh, 2009, p. 1606). To capture the belief that μ_t is stable, they imposed a prior that the predictive R^2 from the regression of r_{t+1} on μ_t is not very large, which is equivalent to the belief that the total variance of μ_t is not very large relative to the variance of realized returns. To capture the belief that μ_t is persistent, they impose a prior that β, the slope of the AR(1) process for μ_t, is smaller than 1 but not by much. Furthermore, they favor sufficiently negative values of ρ_{uw} to reflect the prior that expected returns are countercyclical.

1.2.2 The CIR predictive system

Empirical evidence by Henkel et al. (2011) and others shows that return *predictability* is markedly countercyclical, i.e., it is much stronger during economic recessions. Higher predictability means that the percentage of the variance of realized returns explained by variations in expected returns is higher. Hence, this evidence suggests that the variance of expected returns should increase during recessions and therefore be time-varying.

On the other hand, the conditional variance of the AR(1) process used in the predictive system of Pástor and Stambaugh (2009) to model expected returns is constant over time and therefore excludes by construction a countercyclical degree of predictability. We propose a parsimonious departure from the AR system that does not introduce any additional parameters, but instead modifies the governing equation of the latent expected return process in a way that integrates the aforementioned economically motivated features, as well as additional important implications described below.

Suppose that the true unobservable process μ follows a Cox et al. (1985) (CIR) model,

[7] Other studies on return predictability using state-space models include Conrad and Kaul (1988); Lamoureux and Zhou (1996); Ang and Piazzesi (2003); Brandt and Kang (2004); Duffee (2007); Rytchkov (2012); Piatti and Trojani (2014).

[8] Pástor and Stambaugh (2012) presents an alternative state-space representation of the predictive system where r_t and x_t follow a VAR process with an unobserved additional predictor but as they explain this alternative representation is well suited for exploring the role of predictor imperfection which is not our aim in this paper.

which is a continuous-time mean-reverting process given by the Stochastic Differential Equation (SDE):

$$d\mu_t = \kappa(\theta - \mu_t)dt + \sigma\sqrt{\mu_t}dW_t, \tag{1.6}$$

where the constant κ is the (positive) speed of mean reversion, θ the long-term mean of μ, σ the standard deviation of the diffusion term and W is a standard Brownian motion. For our purposes, it is of primary importance to notice that the diffusion factor, $\sigma\sqrt{\mu_t}$, induces a process with a level-dependent time-varying volatility that increases whenever the level of μ increases. Thus, the μ in the CIR model is an heteroscedastic process whose conditional variance varies in tandem with its level. Interestingly, if the dynamics of μ are countercyclical (i.e., increases during recessions), then its variance as well.

Another interesting difference between the AR and the CIR system, is that the latter is much less likely to produce negative values for expected returns. Indeed, the CIR process rules out negative values of μ if its parameters satisfy the condition $\kappa\theta \geq \frac{\sigma^2}{2}$, along with $\mu_0 > 0$ (see Feller, 1951). This characteristic is in line with some interpretations of standard equilibrium models with risk-averse investors that predict positive expected returns (see for instance Merton, 1980, 1993), as well as with the empirical findings by Campbell and Thompson (2008) and Pettenuzzo et al. (2014) who find that restricting expected return estimates to be positive improves the forecasting power of the predictive regression.

In order to develop the economic intuition and implications, we work with a direct Euler discretization of equation (1.6) under the assumption that the former realization of μ_t is positive, that is:

$$\mu_{t+1} = (1-\beta)E_r + \beta\mu_t + \sqrt{\mu_t}w_{t+1} \quad \text{given that } \mu_t \geq 0, \tag{1.7}$$

where $E_r = \theta$ is the constant long-term mean, $\beta = (1 - \kappa\Delta t)$ is the auto-regressive constant, w is gaussian innovation with variance $\sigma_w^2 = \sigma^2\Delta t$ and Δt is the time step chosen in the discretization, i.e., the elapsed time between t and $t+1$ in the time series notation. We chose the above discretization of the CIR model over the exact distribution of the increments of the CIR process (non-central chi-squared, see Gouriéroux and Jasiak, 2006), to avoid using non-Gaussian copulas to model the joint distribution of the innovations u_t, w_t, v_t. The discretization above allows us to use correlated Gaussian increments in the discretized dynamics as in Brigo and Alfonsi (2005) and similar to Pástor and Stambaugh (2009) AR system, keeping a flexible and tractable model to explore the economic implications of the CIR assumption.

Appendix 1.A and 1.B present the mild technical conditions needed for the discretization (1.7) to have a negligible probability to yield negative values for μ. In Appendix 1.C we derive the expressions of the extended Kalman Filter algorithm for a general function $g(\mu_t)$ on the diffusion term, instead of the particular case $g(.) = (.)^{(1/2)}$ of equation (1.7).[9] Notice that the AR(1) system is nested in the general system derived in the Appendix

[9]The derivation in appendix 1.C includes the case $g(.) = \sqrt{|.|}$, which is well defined for $\mu_t < 0$, unlike equation (1.7). However, for simplicity of exposure we develop the economic intuition of the model in the simpler case of equation (1.7) (refer to appendix 1.A for a detailed discussion).

with $g(.) = (.)^{(0)}$.

Our version of the predictive system is similar to Pástor and Stambaugh (2009)'s AR(1) system (discussed above), but uses the state equation (1.7) for μ instead of equation (1.3), while the realized return equation (1.1) as well as the joint distribution of innovations $[u_t \; w_t]'$ in equation (1.5) are kept the same, i.e., $[u_t \; w_t]'$ is multivariate Gaussian, i.i.d. across t, with zero mean and variance-covariance matrix denoted Σ_{11}.

We refer to this new model as the CIR system throughout the rest of the paper. We do not intend to focus on the selection and the analysis of the predictive power of a particular predictor, thus we concentrate our study on the predictive system without predictors, isolating the effect and the implications of the modified equation for μ. In that sense, our analysis remains valid regardless of the predictors that could be integrated into the system, as our results concern the relation between r and μ and are not dependent on a given predictor. In Appendix 1.C we present the equations of the general system including the possible interaction with external predictors.

In what follows we explore the theoretical implications of the CIR predictive system and perform an empirical analysis using quarterly returns of the value-weighted index of all NYSE, Amex, and Nasdaq stocks in excess of the quarterly return on 1-month T-bills obtained from the Center for Research in Security Prices (CRSP). Following Pástor and Stambaugh (2009) and other studies we begin our sample in 1952-Q1 after the Fed was allowed to pursue an independent monetary policy. Our sample ends in 2012-Q4. In Section 1.2.4 we compare the outputs of the system with a predictive regression using the dividend-price ratio as the predictor variable. The latter is computed as the sum of total dividends paid over the last 12 months divided by the current price, using monthly stock returns with and without dividend on the value-weighted index from CRSP.

1.2.3 Structural differences between the CIR and AR systems

Assuming μ_t follows the dynamics in equation (1.7) has several key implications, including the way in which past realized returns impact changes in expected returns. The implications follow from the mechanics of CIR-type processes[10], in which the diffusion term of μ becomes negligible and its mean reverting strength preponderant whenever μ approaches zero. To see this, notice that from equation (1.7), it follows that,

$$\mu_{t+1} - \mu_t = (1 - \beta)(E_r - \mu_t) + \sqrt{\mu_t} w_{t+1}. \tag{1.8}$$

Hence, whenever μ_t approaches zero, its diffusion term as well, making μ_{t+1} much more likely to increase driven by the mean-reversion term $(1 - \beta)(E_r - \mu_t)$. Notice as well that the speed of mean reversion decreases with the level of the persistence parameter β, thus being two competing effects on the dynamics of expected returns.

Using the Wold representation MA(∞) of the AR(1) process, Pástor and Stambaugh (2009) explore the interesting temporal dependence of returns in the predictive system[11].

[10] Similar processes such as Constant Elasticity of Variance (CEV), considered in the general derivation of the Kalman Filter in Appendix 1.C, present the same type of mechanics.

[11] For the Wold representation of an AR(1) process see for instance Engle and Granger (1987), Anderson (2011),

A similar representation can be done for the CIR-type process applying backward iteration of equation (1.7) and assuming the positivity of the trajectory, which yields:

$$\mu_t = E_r + \sum_{i=0}^{\infty} \beta^i w_{t-i} \sqrt{\mu_{t-1-i}}. \tag{1.9}$$

Equation (1.9) shows that, unlike the AR system, in our setting past innovations on expected returns occurring at times when the level of μ is relatively higher, have a higher impact in future expected returns than innovations occurring at times of low μ, everything else being equal. In other words, the key mechanism is that, unlike with an AR(1), μ_t is now a weighted function of both past expected returns and the corresponding shocks.

As discussed in section 1.2.4, this dependence on the level of μ_t implies variations in the conditional variance of expected returns, and such variance variations are correlated with variations in the level of expected returns. If the change effect prevails over the level effect (as expected), these model predictions are in line with the empirically established countercyclical property of the degree of return predictability. These effects are absent in the AR system, in which the filtered conditional variances of expected returns are constant over time.

Using equations (1.1) and (1.9), the return k periods ahead can be written as:

$$r_{t+k} = E_r + \sum_{i=0}^{\infty} \beta^i w_{t+k-1-i} \sqrt{\mu_{t+k-2-i}} + u_{t+k}. \tag{1.10}$$

Using equation (1.10) it can be shown that the autocovariance of r_t is equal to (cf. Appendix 1.B for the derivation):

$$\mathrm{Cov}(r_{t+k}, r_t) = \beta^{k-1} \left(\underbrace{\beta \sigma_\mu^2}_{\text{level effect}} + \underbrace{\sigma_{uw} \mathbb{E}(\sqrt{\mu_{t-1}})}_{\text{change effect}} \right), \tag{1.11}$$

where σ_μ^2 is the unconditional variance of μ and is given by:

$$\sigma_\mu^2 = \frac{\sigma_w^2 E_r}{1 - \beta^2}. \tag{1.12}$$

In the AR predictive system the expressions for the autocovariance of r_t and the variance of μ are equal to equations (1.11) and (1.12), except that the unconditional expectation terms, $\mathbb{E}(\sqrt{\mu_{t-1}})$ in (1.11) and E_r in (1.12), do not appear. As we will see, equations (1.11) and (1.12) have important implications for the predictive system.

The R^2 of the regression of r_{t+1} on μ_t captures the fraction of variance in r_{t+1} explained by variations in μ_t. Hence it measures the level of predictability of returns and is defined as,

$$R^2 = \frac{\sigma_\mu^2}{\sigma_r^2} = 1 - \frac{\sigma_u^2}{\sigma_r^2}, \tag{1.13}$$

Cochrane (2005) chapter 6.

then, a lower variance for μ implies a lower R^2. From the corresponding expressions for σ_μ for the AR and CIR systems, we have,

$$R^2_{ar} = \frac{\sigma^2_{w,ar}}{\sigma^2_r(1-\beta^2_{ar})}, \tag{1.14}$$

$$R^2_{cir} = \frac{\sigma^2_{w,cir}E_r}{\sigma^2_r(1-\beta^2_{cir})}, \tag{1.15}$$

where the *ar* and *cir* subscripts indicate the system considered. In what follows we discuss two opposite cases to analyze the theoretical implications of the CIR system. First, we keep σ_w, β and ρ_{uw} equal in both systems, which yields differences on R^2 and on the relative importance of the change and level effects. Second, we assume that both models have the same R^2, β and ρ_{uw}, in which case the change and level effects have the same relative weights in both systems and the parameter adjustment to get the same R^2 is done trough σ_w.

On the one hand, notice that equations (1.11) and (1.12) imply that, if $\mathbb{E}(\sqrt{\mu_{t-1}}) < 1$ and $E_r < 1$ (a very mild assumption given historical returns), in the CIR system the variance of expected returns and the autocovariance of returns would be lower than in the AR system, everything else equal, that is, assuming σ_w, β, ρ_{uw} are equal for both models. Intuitively, this result steams from the fact in the CIR system μ_t must vary in a limited space of feasible values, unlike in the AR system for which μ_t can take negative values. Mechanically, we can see that expected returns innovations w are weighted by the previous value of $\sqrt{\mu}$ in equation (1.8) and thus, the size of the random shock in μ_{t+1} is much smaller in the CIR system when μ_t approaches zero, everything else being equal.

Notice that the level and change effects can be "mapped" into the autocovariance of r. For $\rho_{uw} < 0$, whenever $\beta\sigma^2_\mu < -\sigma_{uw}\mathbb{E}(\sqrt{\mu_{t-1}})$ returns have a negative autocovariance, and the change effect prevails. Hence, ρ_{uw} needs to be sufficiently negative for this to happen. Setting σ_w, β, ρ_{uw} in the CIR system equal to the corresponding parameters of the AR system induces a change in the relative weight of the two terms reflecting the change and level effects in equation (1.11) with respect to the AR system, through the $\mathbb{E}(\sqrt{\mu_{t-1}})$ and E_r terms in the variance and autocovariance expressions. Assuming $\mathbb{E}(\sqrt{\mu_{t-1}}) \approx \sqrt{E_r}$, and $E_r < 1$ then $E_r < \sqrt{E_r}$; thus the change effect would have a larger relative weight in the return autocovariance of the CIR system (1.11) relative to AR system, keeping σ_w, ρ_{uw} and β constant (regardless of the level of E_r as long as $E_r < 1$).

Indeed, in the CIR system, the knife-edge value of ρ_{uw}, i.e., the value such that the change and level effects exactly offset each other in (1.11) is:

$$\text{k-e } \rho_{uw} = \frac{-\beta_{cir}\ \sigma_{w,cir}}{\sigma_u(1-\beta^2_{cir})} \times \frac{E_r}{\mathbb{E}(\sqrt{\mu_{t-1}})}. \tag{1.16}$$

This expression is equal to the knife-edge of the AR system, except for the ratio $\frac{E_r}{\mathbb{E}(\sqrt{\mu_{t-1}})}$ (which is absent in the AR system). As a consequence, for equal σ_w, β, ρ_{uw}, the knife-edge value of ρ_{uw} for the CIR system would be closer to zero (i.e., less negative) than for the AR system as $\frac{E_r}{\mathbb{E}(\sqrt{\mu_{t-1}})} < 1$. This would imply a less restrictive condition on the level of

ρ_{uw} for the change effect to prevail.

On the other hand, notice that the autocorrelation at k lags of returns in the CIR model can be written as a function of R^2, β and ρ_{uw},

$$\text{Corr}(r_{t+k}, r_t) = \beta_{cir}^{k-1}\left(\beta_{cir}R_{cir}^2 + \rho_{uw}\sqrt{(1-R_{cir}^2)R_{cir}^2(1-\beta_{cir}^2)}\frac{\mathbb{E}(\sqrt{\mu_{t-1}})}{\sqrt{E_r}}\right), \qquad (1.17)$$

which follows from equation (1.11) and noticing that $\sigma_\mu^2 = R^2\sigma_r^2$, $\sigma_u^2 = (1-R^2)\sigma_r^2$ and $\sigma_{w,cir}^2 = \frac{R_{cir}^2\sigma_r^2(1-\beta_{cir}^2)}{E_r}$. The autocorrelation expression (1.17) is equal to the autocorrelation of the AR system (see Pástor and Stambaugh, 2012, equation 22) except for the ratio $\frac{\mathbb{E}(\sqrt{\mu_{t-1}})}{\sqrt{E_r}}$. Assuming $\mathbb{E}(\sqrt{\mu_{t-1}}) \approx \sqrt{E_r}$, the term cancels out[12] in (1.17), thus if the two systems have equal R^2, β and ρ_{uw} then the level and change effects have the same relative weight in both systems. Another way to see this, is equating (1.14) and (1.15), i.e., if $R_{ar}^2 = R_{cir}^2$ then $\sigma_{w,ar}^2 = \sigma_{w,cir}^2 E_r$ and the extra ratio in the knife-edge formula (1.16) disappears, yielding the same expression in both systems.

Equation (1.15) shows that R_{ar}^2 and R_{cir}^2 are likely to diverge whenever E_r is relatively close to zero, while they should tend to be closer as E_r is further from zero (more on this in section 1.3.1). Thus, the difference in R^2 between the two systems is likely to be more accentuated for higher return frequencies (e.g., E_r is closer to zero for quarterly than for yearly return).

Investors may have priors on R^2 (as in Pástor and Stambaugh, 2009; Kvašňáková, 2013) or β (as in Pástor and Stambaugh, 2009), while intuition for σ_w^2 is less clear and σ_r^2 is almost observable and strictly equal for both models. Thus, in our out-of-sample Bayesian analysis (section 1.3.3) we use the same priors as in Pástor and Stambaugh (2009) for β and ρ_{uw} in both systems, and choose prior distributions for $\sigma_{w,ar}^2$ and $\sigma_{w,cir}^2$ such that the prior distribution of R^2 is equal in both systems. As we will see from the posterior distributions of β (Panels A and B of Figure 1.9), β_{cir} tends to be smaller than β_{ar}, and to steadily decrease over time from 1975 to 2012. Hence, μ_t tends to be less persistent in the CIR system. This result is consistent with the discussion of equation (1.8) above, since μ should have a higher speed of mean reversion (lower β) whenever it approaches zero; an effect absent in the AR system. The result indicates that, while most of the adjustment in parameters to obtain the same R^2 happens through σ_w, some of it is due to differences in β. Notice that the relative weight of the term reflecting the level effect in equation (1.11) increases with the level of β. Thus, the more persistent expected returns are, the more important the level effect is. Conversely, the lower the variance of expected returns, σ_μ^2, the weaker the level effect relative to the change effect, indicating that the change effect can be even more likely to prevail in the CIR system.

Another variation of the widely used AR(1) representation for expected returns has been introduced in Van Binsbergen and Koijen (2011). In their specification, when μ_t gets closer to zero, its value is also more likely to be pushed back to its long-term mean rather

[12] The square root of the sample average (unbiased estimator of the expected value) is a consistent although biased estimator of the expected value of the square root of the sample average (see for instance Barreto and Howland, 2006, p. 396). In our in-sample analysis of section 1.2.4, we find empirically that the ratio $\frac{\mathbb{E}(\sqrt{\mu_{t-1}})}{\sqrt{E_r}}$ ranges between 0.99 and 1.00 for the R^2 and ρ_{uw} considered.

than decrease, as in the CIR system. In Van Binsbergen and Koijen (2011), this effect is due to the time-varying autoregressive parameter but the conditional heteroscedasticity of μ introduced in our predictive system is not present in their model.

The state-space representation of the system illustrates other angles of structural features of the CIR model. Using the Kalman filter, the conditional expected return can be written as the unconditional expected return mean plus linear combinations of past return forecast errors[13], where forecast error for the return in each period is defined as $\epsilon_t = r_t - \mathbb{E}(\mu_t|D_{t-1})$. Then the expected return conditional on the history of returns can be written as

$$\mathbb{E}(\mu_t|D_t) = E_r + \sum_{s=0}^{\infty} \lambda_s \epsilon_{t-s}, \tag{1.18}$$

where $\lambda_s = \beta^s m$, and m is a steady-state filter parameter. Similar to the AR(1) predictive system of Pástor and Stambaugh (2009), whenever the change effect is greater than the level effect, i.e., the covariance term in the autocovariance (1.11) is larger than $\beta\sigma_\mu^2$, then $m < 0$ for sufficiently negative ρ_{uw} (see Appendix 1.C).

The conditional expected return can also be written as a function of past returns instead of past forecast errors as follows:

$$\mathbb{E}(\mu_t|D_t) = E_r + \sum_{s=0}^{\infty} \omega_s (r_{t-s} - E_r), \tag{1.19}$$

where, in steady state,

$$\omega_s = (\beta - m)^s m. \tag{1.20}$$

Equations (1.18) and (1.19) have the same structure than the equivalent expression in the AR predictive system, but the coefficient m is modified, leading to different predictions, induced by a more time-varying behavior of the λ and ω coefficients. The expressions for the finite sample value of m_t (equation 1.60) and the steady-sate value of m (equation 1.72) in the CIR system are derived in the Appendix 1.C. We refer the reader to the appendix in Pástor and Stambaugh (2009) to see the corresponding expressions in the AR system.

The conditional expected return depends on the true unconditional mean, E_r. Using the sample mean to estimate E_r in equation (1.19) and truncating the summation in the right-hand side to $s = t - 1$, yields an estimate of the expected return $\mathbb{E}(\mu_t|D_t)$ equal to a weighted average of past returns, i.e.,

$$\mathbb{E}(\mu_t|D_t) = \sum_{s=0}^{t-1} \kappa_s r_{t-s}, \tag{1.21}$$

[13] In the case where predictors are used, an additional term containing innovations in the predictors is added to equations (1.18), (1.19) and (1.21), for details, see Appendix 1.C.

where

$$\kappa_s = \frac{1}{t}\left(1 - \sum_{l=0}^{t-1} \omega_l\right) + \omega_s, \tag{1.22}$$

and $\sum_{s=0}^{t-1} \kappa_s = 1$. This expression has the same form than in the AR system. However, the ω_s are functions of m, which has a different expression in the CIR system that depend on the values of μ (See equation (1.72) of Appendix 1.C).

In order to illustrate the implications of the differences between the CIR and AR models, such as the differences in the coefficients of the model representations above, in what follows we perform an in-sample comparison of the two systems for the same stock market series. In section II.B, Pástor and Stambaugh (2009) tested different values for ρ_{uw} to see its impact on the predictive system outputs. Their analysis is performed assuming $\beta = 0.9$, $R^2 = 5\%$ and using the sample estimates for σ_r and E_r in the system without predictors. In that case, all other parameters needed to estimate μ are functions of the parameters mentioned. The last row of Table 1.1 presents the numerical values of the parameters used in Pástor and Stambaugh (2009) section II.B, adjusted using the updated sample estimates for E_r and σ_r. In the Internet Appendix we explore the implications for the CIR system of setting $\sigma_{w,cir} = \sigma_{w,ar}$ (leading to different R^2 for each system), which corresponds to the first row in Table 1.1 for the CIR system and the last row for the AR system. As shown in the table, keeping the same value of σ_w yields an R^2 of 5% and 0.09% for the AR and CIR systems respectively. However, it is common to have prior estimates about the level of R^2, while we do not have prior beliefs about the level of σ_w. Thus we focus the analysis here for the two systems with the same level of $R^2 = 5\%$ and maintain β, σ_r and E_r equal (the resulting parameters of both systems are reported in the last row of Table 1.1). As mentioned above, if R^2 and β are equal in both models, then the knife-edge value of ρ_{uw} is the same in both systems as well (using the approximation $\mathbb{E}(\sqrt{\mu_{t-1}}) \approx \sqrt{E_r}$), equal here to -0.47.

We find that, while the finite-sample estimates of m_t converge fast to the steady-state value in the AR system, this is not the case for the CIR system. In effect, given its dependence on the current level of μ, the finite-sample values of m_t present significant variations over time around the steady-state value. This feature has important implications for the conditional variance of μ which also implies a time-varying predictability of returns.

For example the changing dynamics of μ in the CIR system can be mapped in the dynamics of the coefficients connecting the realized returns and expected return estimates of the system. As Figure 1.1 shows for several values of ρ_{uw} (below or equal to the knife-hedge value), unlike the AR system, the CIR system has a *conditional memory* (conditional on the phase of the market) reflected for instance in the κ coefficients, capturing the potential long-lasting effects of past shocks on future expected returns. In effect, the lagged κ's and similarly λ, ω (unreported for space consideration) of the CIR system are different, whether they are calculated after a series of positive realized returns or after a series of negative realized returns. In order to illustrate the conditional memory of the CIR system we zoom in two particular sub-periods in the sample, i.e., 1952Q1 to 1999Q4 (end of the bull market of the IT bubble), and 1952Q1 to 2002Q4 (end of the bear

market after the IT bubble bursts). Figure 1.1 presents the corresponding values of the finite-sample coefficients κ (the behavior of λ and ω coefficients is similar), for both systems calculated at the end of the two sub-periods. From the figure we see that, when the change effect dominates, in the CIR system the latest observations have a larger impact in expected returns following a falling market than after a rising market, that is, expected return estimates depend more on the most recent returns during bad times (through a more negative relation with the latter). This effect is induced by the proximity of μ_t to its lower bound (zero) after a long bull market, as further positive returns cannot push μ_t much further down.

In Figure 1.1 we observe a remarkable difference in the shape of the corresponding coefficients curves for the two systems. Indeed, the rise from the most negative values of the coefficients corresponding to the lower lags (i.e., most recent values) toward zero of the curve of κ, is much sharper in the CIR system. This implies that the latest return will induce a larger (relative) correction of the estimate of μ_t in the CIR system relative to the AR system. To see this, consider any two lag indices l_1 and l_2 such that $0 < l_1 < l_2 < 20$ (notice that the lower lag l_1 refers to a more recent observation than l_2) in Figure 1.1, and remark that[14] $\frac{|\kappa_{l_1}^{cir}|}{|\kappa_{l_2}^{cir}|} > \frac{|\kappa_{l_1}^{ar}|}{|\kappa_{l_2}^{ar}|}$. The coefficients of the lower lags are also larger in absolute terms for the CIR system, i.e., $|\kappa_1^{cir}| > |\kappa_1^{ar}|$. Thus, the CIR system gives more (negative) weight to the latest observations than the AR system. This effect is due to a higher time dependence of m_t with respect to the corresponding terms in the AR system (cf. Appendix 1.C), and it is also observed in the weights calculated at the end of the total sample, 2012Q4 (unreported for space considerations).

Panels A and B of Figure 1.2 present the filtered equity premium, $\mathbb{E}(r_{t+1}|D_t)$, for the two systems on the overall sample period when their R^2 is set to 5%. We observe in particular that the level of μ reaches larger maximum values in the CIR system (Panel B) than in the AR system (Panel A), while it tends to slow its variations much faster as it approaches zero as suggested by the dynamics of the CIR model.

1.2.4 Economic Implications

Remarkable implications of using the CIR-discretization to model expected returns arise from the relationship it implies between the level of expected return and important quantities such as the conditional covariance between expected and realized returns, the conditional autocovariance of returns, and ultimately the conditional degree of return predictability.

Following Campbell (1991), the unexpected return can be decomposed as the sum of a cash flow shock $n_{C,t+1}$, and a discount rate shock $n_{E,t+1}$:

$$u_{t+1} = n_{C,t+1} - n_{E,t+1}. \tag{1.23}$$

Under the CIR assumption for the dynamics of expected return μ_t, the discount rate shock

[14] We also observe $\frac{|\lambda_{l_1}^{cir}|}{|\lambda_{l_2}^{cir}|} > \frac{|\lambda_{l_1}^{ar}|}{|\lambda_{l_2}^{ar}|}$ and $\frac{|\omega_{l_1}^{cir}|}{|\omega_{l_2}^{cir}|} > \frac{|\omega_{l_1}^{ar}|}{|\omega_{l_2}^{ar}|}$, unreported here.

is equal to[15] $\sqrt{\mu_t}w_{t+1}$. This is a key observation which implies the following conditional properties of the CIR system:

$$\text{Var}(n_{E,t+1}|D_t^*) = \text{Var}(\mu_{t+1}|D_t^*) = \mu_t\sigma_w^2, \tag{1.24}$$

$$\text{Cov}(r_{t+1}, \mu_{t+1}|D_t^*) = \text{Cov}(u_{t+1}, \sqrt{\mu_t}w_{t+1}|D_t^*) = \sqrt{\mu_t}\sigma_u\sigma_w\rho_{uw} \tag{1.25}$$

$$\text{Cov}(r_{t+1}, r_{t+k}|D_t^*) = \text{Cov}(u_{t+1}, \mu_{t+k}|D_t^*) = \beta^{k-2}\sigma_u\sigma_w\rho_{uw}\sqrt{\mu_t}. \tag{1.26}$$

where $D_t^* = \{r_t, \mu_t, r_{t-1}, \mu_{t-1}, ..., r_1, \mu_1\}$ represents the history of realized and expected returns[16] at time t.

First, equation (1.24), shows that in the CIR model the conditional variance of expected returns is time-varying and proportional to the level of μ_t. Hence, during bad times, when expected returns are higher (i.e., assuming ρ_{uw} is sufficiently negative), the variance of the discount rate shock is higher. Second, equation (1.25) shows that the strength of the relationship between realized and expected return is time-varying, depending on the level of μ_t. For $\rho_{uw} < 0$, it implies an empirically plausible negative time-varying conditional covariance between realized and expected returns, that is higher (in absolute terms) when μ_t is high, i.e., during bad times. Third, the conditional auto-covariance of returns (1.26) is time-varying, exponentially decreasing and proportional in absolute value to $\sqrt{\mu_t}$. It follows that returns are more persistent during bad times. These last three implications all imply a higher return predictability during recessions. This feature of the model is in line with the economic intuition that investors demand higher risk premiums in bad times, when volatility is higher as well, as "*overall adjustments to discount rates per unit of change in economic state are larger in bad times*", Henkel et al. (2011).

Henkel et al. (2011) also notice that price-dividend ratios become more volatile and prices more sensitive to changing expectations as conditions worsen. The fact that the variance of the dividend yield is larger during bad economic times is quite intuitive. As noted by Goetzmann and Jorion (1993), given the persistence of dividend payments, the variations in dividend yields are essentially driven by price changes, which are more volatile during recessions. At the same time, return predictability and variations of the dividend yield are tightly linked (see Cochrane, 2008).

Henkel et al. (2011) find empirically that the predictive power of the dividend yield is significant during economic contractions but nonexistent during expansions. In order to illustrate this point empirically with our model, in Panel A of Figure 1.3 we present the filtered conditional volatility of expected returns for the CIR system (without predictors) in finite-sample, with a prior for ρ_{uw} set to -0.85 (the conditional volatility of the AR system is also presented). Moreover, Panel A of Figure 1.3 also presents the conditional variance of fitted values from a standard OLS predictive regression using a rolling window with the latest 30 years of data (as in Fama and French, 1989) using the dividend-price ratio

[15] This result can be derived using the Wold-type representation of the CIR process and well as the decomposition of Campbell (1991).

[16] In practice, the past values of expected returns are filtered using the Kalman filter algorithm presented in Appendix 1.C.

as the predictor variable x. For the regression, the conditional variance of μ is calculated as the squared slope b^2, multiplied by the variance of x (i.e., the dividend-price ratio), both estimated using the latest 30 years of data. We observe that the conditional variance of μ for the CIR system increases significantly during economics recessions (grayed areas in the figure) as reported by the National Bureau of Economics Research. Interestingly, similar increases are observed for the predictive regression using the dividend-price ratio, while the conditional variance of μ in the AR system is constant over time.

The degree of return predictability conditional on the information set D^*_{t-1} can be measured as the fraction of the variance in returns explained by variations in expected returns for the CIR model as:

$$R^2(r_{t+1}, \mu_t | D^*_{t-1}) = \frac{\mathrm{Var}(\mu_t | D^*_{t-1})}{\mathrm{Var}(r_{t+1} | D^*_{t-1})} = \frac{\mu_{t-1}\sigma_w^2}{\mu_{t-1}\sigma_w^2 + \sigma_u^2} = \frac{1}{1 + \frac{\sigma_u^2}{\mu_{t-1}\sigma_w^2}}. \tag{1.27}$$

Together with assumption $\rho_{uw} < 0$, equation (1.27) yields an empirically plausible degree of predictability that concentrates in bad times, when realized returns are low and persistent and when at the same time expected returns are high and highly volatile. This is in line with Henkel et al. (2011) who find that the conditional R^2 of a predictive Regime-Switching Vector Autoregressive model presents significant variations across the business cycle, using several predictors including the dividend yield. In order to illustrate this point empirically, we compute a time-conditional R^2 in our rolling predictive regression using the dividend yield:

$$R^2(t+1|D_t) = \frac{\sigma^2_\mu(t|D_t)}{\sigma^2_r(t+1|D_t)}, \tag{1.28}$$

where $\sigma^2_\mu(t|D_t)$ corresponding to the conditional variance of μ, is calculated as explained above, and $\sigma^2_r(t+1|D_t)$, the conditional variance of returns, is estimated using the sample estimate for the variance of r with the latest 30 years of data. Notice that the term (1.28) is similar in spirit to to the conditional R^2 considered in Henkel et al. (2011).

Similarly, we calculate the conditional R^2 of the CIR (and AR) predictive system as given by the right-hand side of (1.28), using the corresponding filtered values for $\sigma^2_\mu(t|D_t)$ and $\sigma^2_r(t+1|D_t)$ which correspond to the quantities Q_t and the first component of S_{t+1} respectively, for which expressions are provided in the Kalman Filter derivation of the system in Appendix 1.C. Panel B of Figure 1.3 presents the conditional R^2 of equation (1.28), i.e., the ratio of the conditional variances of μ_t and r_{t+1} for both systems and for the (rolling) predictive regression using the dividend yield.

The time series of the conditional R^2 of the CIR system without predictors and the corresponding series of the predictive regression present remarkably similar dynamics, increasing and decreasing over time in tandem. Moreover, the conditional R^2 increases during economic recessions. We find a qualitatively equivalent result when using the same conditional variance estimate of returns used in the predictive regression instead of the first component of S_{t+1} in the formula of the conditional R^2 of the CIR system (unreported). This result confirms the counter-cyclicality of the predictability of the dividend yield documented by Henkel et al. (2011). More importantly, it shows that the CIR predictive

system *without predictors* naturally reproduces the countercyclical dynamics of return predictability (an effect absent in the AR system which presents a constant conditional R^2 as shown in Figure 1.3 Panel B). Interestingly, in both panels of Figure 1.3, we note that after the year 2000, the curves corresponding to the CIR system and dividend-yield regression start to diverge significantly in terms of level. Goyal and Welch (2003) highlight an instability of dividend price ratio autoregression coefficient, which has increased from about 0.4 in 1945 to about 0.9 in 2000 according to their estimation procedure. This observation may explain the behavior of the predictive regression curves in Figure 1.3 as an increasing autoregression coefficient implies a lower conditional variance in dividend yield and thus a lower conditional R^2. This post-2000 effect is absent in the CIR system as the autoregressive parameter used in this illustration is constant over time.

Finally, the CIR system implies a natural risk-return relation that holds both for unconditional an conditional estimates of expected returns and volatility. To see this, from equation (1.12), it follows that the unconditional variance of realized returns is

$$\sigma_r^2 = \sigma_\mu^2 + \sigma_u^2 = \frac{\sigma_w^2 E_r}{1-\beta^2} + \sigma_u^2 \tag{1.29}$$

thus there is a positive relationship between the unconditional variance of realized returns and the unconditional level of expected returns in the CIR model, which is absent in the AR system. In other words, the CIR implies an *unconditional* positive risk-return tradeoff.

In order to test whether there is a relationship between the conditional expected returns and the conditional volatility of realized returns according to the systems, we regress the conditional estimates for expected returns and volatility from the Kalman filter for the CIR and AR models, as follows:

$$\mathbb{E}(r_{t+1}|D_t) = a + b\sqrt{\mathrm{Var}(r_{t+1}|D_t)} + e_t \tag{1.30}$$

where $\mathbb{E}(r_{t+1}|D_t) = \mathbb{E}(\mu_t|D_t)$ is the filtered expected return from the predictive systems, and $\mathrm{Var}(r_{t+1}|D_t)$ is the expected variance of returns estimated by the systems. Table 1.2 presents the result of the regressions (1.30) for three levels of ρ_{uw}, which correspond to the knife-edge value and two values consistent with the prior that expected returns are counter-cyclical. For the knife-edge value of ρ_{uw}, the change and level effects cancel each other, yielding a constant μ and thus no relation with the conditional variance in both systems. On the other hand, for the other values of ρ_{uw}, the results show that the conditional estimates of the CIR model have a positive relationship, while the estimates from the AR have a negative and insignificant relationship, confirming the result in equation (1.29).

1.3 Empirical analysis

In this section we present an out-of-sample analysis of return predictability using both the predictive system and the predictive regression on the dividend yield. We employ a

Bayesian approach for estimating the predictive system and ordinary least squares (OLS) for estimating the regression. First, we discuss the implications of using the CIR system, on the priors of the parameters distributions. Then we conduct an exploratory analysis using point estimate parameters for implementing the systems. Finally, we present our Bayesian analysis and out-of-sample results.

1.3.1 What are plausible values for R^2, β and E_r?

The empirical implementation and estimation of the predictive system in a Bayesian setting needs a set of priors on the distribution of the parameters involved. These priors should represent a plausible parameter set, compatible with the hypotheses behind the system. Unlike the AR system, one of the implications of using the CIR system is the assumption that the expected return process is unlikely to be negative. Hence, the priors used to implement the CIR system should be compatible with this hypothesis. In this section we discuss the plausible values for the R^2 of the regression of r_{t+1} on μ_t, the persistence parameter β and the long-term mean E_r using the same stock market index and sample period as in Pástor and Stambaugh (2009).

The CIR continuous time model ensures the non-negativity of the process μ in equation (1.6) if the parameters respect the condition $\kappa\theta \geq \frac{\sigma^2}{2}$ (Feller, 1951), which can be translated in terms of the parameters of the discretized process as :

$$(1-\beta)E_r \geq \frac{\sigma_w^2}{2}. \tag{1.31}$$

By definition of σ_μ, condition (1.31) also implies an upper bound for σ_μ^2 in the CIR model:

$$\sigma_\mu^2 = \frac{\sigma_w^2 E_r}{1-\beta^2} \leq \frac{2(1-\beta)E_r^2}{1-\beta^2}. \tag{1.32}$$

Furthermore, using an estimate for the variance of realized returns σ_r^2, this condition also provides an upper bound for the R^2 from the regression of r_{t+1} on μ_t since by definition (1.13) it follows that,

$$R^2 = \frac{\sigma_\mu^2}{\sigma_r^2} \leq \frac{2(1-\beta)E_r^2}{(1-\beta^2)\sigma_r^2}. \tag{1.33}$$

This means that, for a given set of plausible values for E_r and β, the CIR positivity condition (1.31) restrains the possible value set for σ_μ, and for R^2 for a given estimate of σ_r. On the other hand, there is no such internal coherence restriction in the AR system as there is no positivity constraint. From condition (1.31) and equation (1.33), it follows that to estimate the upper bound of R^2 for the CIR system, we need to get: 1) E_r^*: the highest plausible value of E_r, and 2) β^*: the lowest value for β. Considering the return sample used in Pástor and Stambaugh (2009), to estimate the highest plausible value of E_r, we calculate $\max(\hat{E}_r(1,..s_0))$ for a sample of size s_0 for $s_0 = \{80, 81, ...208\}$, where 80 points corresponds to a minimum sample period of 20 years and 208 is the total sample size of quarterly data from 1952-Q1 to 2003-Q4 used in Pástor and Stambaugh (2009), which yields $E_r^* = \max(\hat{E}_r(1,..s_0)) = 0.0226$ (9% p.a.). Second, we use the 5% quantile

of the prior distribution of β in Pástor and Stambaugh (2009), i.e., $\beta \sim \mathcal{N}(0.99, 0.15^2)$, hence $\beta^* = 0.99 + 0.15 \times z_{(5\%)} = 0.743$, where $z_{(5\%)}$ denotes the 5% quantile of a random variable with standard normal distribution. These two estimates for E_r^* and β^* within condition (1.31) together with the sample estimate of $\hat{\sigma}_r = 0.0837$ yields an upper bound for the R^2 in the CIR system of 8.37%. This upper bound is in line with Pástor and Stambaugh (2009)'s statement that a plausible prior distribution of R^2 would have most of its mass being below 5% and a mode around 1% for the US stock market quarterly data sample considered.

Note that, using the knife-edge formula of ρ_{uw}, it is possible to derive bounds for the persistence parameter β in order to have a plausible knife-edge value, for both, the AR and CIR systems for a given value of R^2. Considering a knife-edge value within the interval [-1, 0] is equivalent to not excluding the possibility that the change effect dominates over the level effect. Notice that, if $\rho_{uw} > 0$ there is no change effect at all. If the knife-edge value of ρ_{uw} of the AR system is between -1 and 0 then,

$$|\text{k-e } \rho_{uw}| = \left| \frac{-\beta \sigma_w}{\sigma_u (1 - \beta^2)} \right| < 1. \tag{1.34}$$

From the definition of R^2 and σ_μ for the AR system, notice that, $\sigma_w^2 = (1 - \beta^2) R^2 \sigma_r^2$ and $\sigma_u^2 = (1 - R^2)\sigma_r^2$. Replacing these in inequality (1.34) and squaring yields

$$\begin{aligned} \frac{R^2 \beta^2}{(1 - R^2)(1 - \beta^2)} &< 1 \\ |\beta| &< \sqrt{\frac{C}{1 + C}} \end{aligned} \tag{1.35}$$

where $C = \frac{(1 - R^2)}{R^2}$, which simplifies to $|\beta| < \sqrt{1 - R^2}$. For instance, if $R^2 = 5\%$ the upper bound of β is 0.97. A similar calculation using the knife-edge value for the CIR system (equation 1.16) yields the bounds for β in the CIR model which are given by (1.35) but with $C = \frac{(1 - R^2)}{R^2 E_r}$. The consistency condition (1.35) implies an inverse relationship between R^2 and the maximum feasible persistence parameter of μ. If one believes that μ is very persistent and we do not preclude the possibility that the change effect dominates over the level effect then the R^2 cannot be very high and viceversa.

The positivity condition (1.31) implies a long-term mean strictly positive, i.e., $E_r > 0$. Pástor and Stambaugh (2009)'s prior distribution for E_r is Gaussian with a "large" 1% standard deviation around its sample mean, denoted $\hat{E}_r$ (see p. 9 of the internet appendix in Pástor and Stambaugh, 2009). In effect, the 1% quantile of such distribution is a negative number, which is incompatible with the CIR model assumption, especially if we consider that the presumably positive process μ should vary around E_r. Indeed, assuming that expected returns are non-negative implies that the long-run average of μ, E_r should be "far enough" from zero. Thus, a plausible value for the variance of the prior distribution of E_r should be lower than 1% for the CIR system. For instance, given that the prior for the distribution of E_r is symmetric, one may assume that the distance between the sample mean $\hat{E}_r = 0.0185$ and its highest plausible value of $E_r^* = 0.0226$ (see calculation

above), is the same distance between $\hat{E}_r$ and a low quantile of its distribution. We deduce σ_{E_r} as follows. Assume $\hat{E}_r + \sigma_{E_r} z_{(5\%)} = \hat{E}_r - (E_r^* - \hat{E}_r)$, hence

$$\sigma_{E_r} = \frac{(\hat{E}_r - E_r^*)}{z_{(5\%)}},$$

which is 0.25% for the sample estimates mentioned above, thus four times smaller than the prior standard deviation of 1% used in Pástor and Stambaugh (2009) for the AR model. This choice ensures a positive 1% quantile for the prior distribution of E_r.

1.3.2 Out-of-sample return prediction using point estimate parameters

In what follows we present the results of an exploratory analysis which consists in an out-of-sample return prediction exercise of quarterly returns of the value-weighted CRSP US aggregate stock market index in excess of the quarterly return on 1-month T-bills obtained from the Center for Research in Security Prices (CRSP). We begin our sample in 1952, as in our in-sample analysis of section 1.2.4. In order to address concerns regarding the dependence of predictability evidence on the oil price shocks period $1973 - 1974$, we set our out-of-sample period to 1975-2012 for this first analysis. We also consider additional out-of-sample periods in the Bayesian analysis of section 1.3.3.

In order to estimate μ, we use both the CIR and AR predictive systems without predictors, with fixed values for the models parameters. Every quarter we estimate E_r and σ_r using the prevailing sample estimates at each point in time and we set the R^2 of the regression of r_{t+1} on μ_t equal in both models, with $R^2 = 0.5\%, 1\%, 2\%, 3\%, 4\%, 5\%$. We consider two plausible values for the persistence parameter of $\beta = 0.9$ and then $\beta = 0.8$ (kept constant throughout the entire period). Furthermore, we perform predictions using several values of ρ_{uw}, i.e., from -0.95 to 0.95 with a step of 0.05. All other parameter values in the systems follow from their corresponding definitions. We also estimate μ using the OLS predictive regression on the dividend yield. The latter is defined as in section 1.2.4. The regression is re-estimated every quarter using available data. We compute predictions at each point in time using the updated parameters.

Following former studies such as Goyal and Welch (2008), we assess out-of-sample predictive power with the out-of-sample (OS) R^2_{OS} introduced by Campbell and Thompson (2008),

$$R^2_{OS} = 1 - \frac{MSE_{pred}}{MSE_{mean}}, \tag{1.36}$$

where MSE_{pred} is the mean squared error of the model predictions and MSE_{mean} is the mean squared error of using the prevailing return's historical average as estimate of expected return. This metric evaluates weather a given system produces more accurate predictions than the no-predictability random walk benchmark, i.e., the prevailing historical average.

The R^2_{OS} for the predictive regression is negative and is equal to -0.012, i.e., the regression does not produce better predictions than the historical average. The overall results for the predictive systems are presented in Figure 1.4 for $\beta = 0.9$, and Figure 1.5 for

$\beta = 0.8$. The two figures present the R^2_{OS} for each system, as a function of ρ_{uw}, when setting the R^2 of the regression of r_{t+1} on μ_t to a given value in $\{0.5\%, 1\%, 2\%, 3\%, 4\%, 5\%\}$. The grayed areas in both figures correspond to values of ρ_{uw} implying a dominant change effect.

First, we observe that the highest R^2_{OS} for each system are positive, i.e., both predictive systems manage to produce better forecasts than the historical average, as well as the regression on the dividend yield. For each system, the larger R^2_{OS} are obtained with different priors on R^2 and β. The highest R^2_{OS} across all parameter combinations considered corresponds to the CIR system (3.39%), with $\beta = 0.8$ and $R^2 = 2\%$. On the other hand, the highest R^2_{OS} for the AR system is 3.00% and corresponds to $\beta = 0.9$, and $R^2 = 4\%$ (where $\rho_{uw} = -0.95$ in both cases). This result is consistent with the intuition that the CIR system is more suited for lower levels of R^2 (as μ would have a lower variance) and thus with a less a stringent condition on ρ_{uw} to yield a dominant change effect. Furthermore, it also suggests that the CIR system may imply an expected return process with lower persistence, a result confirmed by the posterior distribution of β of the Bayesian analysis of the next section (1.3.3).

Also, we note that for most values considered for the R^2 of the regression of r_{t+1} on μ_t, for both AR and CIR systems, the R^2_{OS} are positive when ρ_{uw} is below its knife-edge value (except the case $\beta = 0.8$; $R^2 = 4\%, 5\%$; $\rho_{uw} = -0.95$ for the CIR system). Table 1.3 presents a summary of the highest and lowest R^2_{OS} obtained for each system as well as the corresponding parameters set. For both systems, the highest R^2_{OS} are obtained with $\rho_{uw} = -0.95$, whereas the lowest R^2_{OS} are obtained for $\rho_{uw} = 0.95$. This result indicates that Pástor and Stambaugh (2009)'s believe that the change effect should dominate, is also consistent with the CIR system. It is also interesting to notice that in all configurations with $\rho_{uw} \geq 0$ (only level effect) the less negative R^2_{OS} are obtained with the CIR system for all priors of R^2 considered.

In the next section, we explore the performance of the predictive system, considering different out-of-sample periods and different priors on the parameters. The results presented thereafter confirm the conclusions above.

1.3.3 Out-of-sample return prediction using Bayesian parameters estimates

Using the same data as in section 1.3.2, we conduct now an out-of-sample analysis using both predictive systems estimated with a Bayesian methodology, as well as the OLS regression on the dividend yield. As in section 1.3.2, the predictive regression is re-estimated every quarter using available data and predictions are calculated at each point in time using the updated parameters. The Bayesian parameter estimation procedure is similar to that of Pástor and Stambaugh (2009) for estimating the predictive systems. This procedure allows incorporating parameter uncertainty and specifying less informative prior distributions. Posterior distributions for the parameters are obtained using Gibbs sampling (see for instance Kim and Nelson, 1999). Following Pástor and Stambaugh (2009), we estimate the predictive systems parameters by simulating 76000 posterior draws, dropping the first 1000 as a "burn-in" period and take every third draw from the rest to obtain

25000 posterior draws. The overall Markov Chain Monte Carlo (MCMC) procedure and the posterior distributions for the AR system are described in the internet appendix of Pástor and Stambaugh (2009). We refer the reader to Appendix 1.D for further details on the Bayesian parameter estimation procedure and the posterior distributions in the CIR system.

The predictive systems parameters are re-estimated on the first available date of each year in the sample, while predictions (estimates of μ) are computed each quarter using the data available at each point in time (thus running the filter with the same parameters over the year).

For the AR system, the priors distributions used are identical (except for E_r as explained below) to those described in section B.5 of the internet appendix of Pástor and Stambaugh (2009). We thus refer to their initial paper for more details and provide only a brief description of the distributions. The prior for β, plotted in Panel B of Figure 1.7, is chosen to capture the belief that μ is persistent, i.e., β is smaller than one but not by much[17] $\beta \sim \mathcal{N}(0.99, 0.15^2) \times \mathbb{I}_{\beta\in(0,1)}$. The prior on E_r is slightly modified in order to use the same prior as the CIR system[18]: $E_r \sim \mathcal{N}(\bar{r}, \sigma^2_{E_r})$, where $\bar{r}$ denotes the mean of the returns $\{r_t\}$ available at the date of estimation and σ_{E_r} is choosen as described in section 1.3.1. We consider three prior distributions for σ_w, plotted in Panel B, D and F of Figure 1.6. The submatrix $\Sigma_{11} = \begin{bmatrix} \sigma_u^2 & \sigma_{uw} \\ \sigma_{wu} & \sigma_w^2 \end{bmatrix}$ has, for each prior, an inverted Wishart distribution: $\Sigma_{11} \sim IW(T_0\hat{\Sigma}_{11,0}, T_0)$, where T_0 is equal to one fifth of the available return sample size. The prior mean $\mathbb{E}(\Sigma_{11})$ is set according to: 1) a prior value $\bar{R}^2$ for its diagonal elements, and 2) our priors on ρ_{uw} (see below) for the non-diagonal elements[19]. The three different priors on σ_w, corresponding to different priors on the R^2 (and the variance of μ), are obtained by setting $\mathbb{E}(\sigma^2_\mu)$ equal to a given percentage of the prevailing sample return variance $\hat{\sigma_r}^2$, i.e., $\mathbb{E}(\sigma^2_\mu) = \bar{R}^2\hat{\sigma_r}^2$ and $\mathbb{E}(\sigma^2_u) = (1-\bar{R}^2)\hat{\sigma_r}^2$ for $\bar{R}^2$ equal to 2.5% (less predictability prior), 5% (prior used in Pástor and Stambaugh, 2009, denoted hereafter benchmark prior) and 10% (more predictability prior). The corresponding prior distributions of the R^2 are presented respectively in Panel A, C and E of Figure 1.6. Moreover, we consider two priors on ρ_{uw} used in Pástor and Stambaugh (2009), which are presented in Panel A of Figure 1.7: noninformative (flat between -0.9 and 0.9) and more informative prior (most of the mass below -0.71).

The priors used for the CIR system are the same for ρ_{uw}, β and R^2. The latter implies a prior distribution with higher levels of σ_w as shown in Panels B, D and F of Figure 1.6 and explained in section 1.2.3. As described in detail in section 1.3.1, the prior distribution of E_r used here for both the CIR and AR systems, is slightly different from the prior used in Pástor and Stambaugh (2009); thought we use the same method to estimate the mean, we use a lower variance, in order to preclude negative value draws for the long-term mean

[17] $\mathbb{I}_{\beta\in(0,1)}$ denotes here the indicator function equal to 1 if $\beta \in (0,1)$ and 0 otherwise. In our case, this corresponds to retain only draws of β satisfying the condition.

[18] We also tried to use the initial prior for E_r used in Pástor and Stambaugh (2009) for the AR system, i.e., $E_r \sim \mathcal{N}(\bar{r}, 0.01^2)$, but the out-of-sample results were poorer, i.e., negative R^2_{OS} for the AR system.

[19] We refer to page 9 of the internet Appendix of Pástor and Stambaugh (2009) for a description of the prior draw procedure of the non-diagonal elements.

of expected returns E_r.

Again, we use the R^2_{OS}, described in section 1.3.2, to assess the out-of-sample predictability of the systems[20] and the regression. The statistical significance of R^2_{OS} is assessed using the F_{MSE} statistic proposed by McCracken (2007), which tests for equal MSE of the historical mean and system's conditional forecasts. It is given by:

$$F_{MSE} = (T - s_0)\frac{(MSE_{mean} - MSE_{pred})}{MSE_{pred}}, \tag{1.37}$$

where T stands for the total size of the sample periods and s_0 for the initial calibration sample. In our tables we use $*$, $**$ and $***$ to indicate statistical significance at 10%, 5% and 1% levels respectively.

In addition to this asymptotic testing approach, we assess the systems' prediction performance using the log predictive score LPS. This metric is calculated as:

$$LPS = \frac{1}{(T - s_0)} \sum_{t=s_0}^{T-1} \log p(r_{t+1}|D^*_t), \tag{1.38}$$

where $p(r_{t+1}|D^*_t)$ is the predictive density of r_{t+1} given realized and (estimated) expected returns available at time t. The latter is available in closed form for both predictive systems and predictive regression, since each return model is conditionally Gaussian $\mathcal{N}(\mu_t, \sigma^2_r)$.

The out-of-sample predictive performance statistics R^2_{OS} and the LPS are presented in Tables 1.4 and 1.5 for the two predictive systems respectively. The results of the regression are given in Table 1.6. We consider four out-of-sample periods. The first one is the longest and is the same period considered in section 1.3.2 with point estimates: 1975-2012. We also consider three additional out-of-sample periods of 25 years, with starting dates spaced by 5 years: 1975-2000, 1980-2005, 1985-2010.

There are four clear conclusions from the results regarding the predictive system, 1) the more informative prior on ρ_{uw} that corresponds to a dominant change effect improves out-of-sample return forecasts for both systems, 2) for all periods and predictability priors considered, using the more informative prior on ρ_{uw}, the CIR system yields better out-of-sample estimates than the AR system, 3) in both systems the prior on the predictability level (i.e. R^2 of the regression of r_{t+1} on μ_t) that yielded the highest R^2_{OS} and LPS numbers was different in each of the out-of-sample periods considered, which suggests a varying level of predictability, 4) the CIR system using the more informative prior on ρ_{uw} yields significantly better out-of-sample predictions than the prevailing historical average for all out-of-sample periods considered, for at least one R^2 prior. In fact, the R^2_{OS} is significant in 10 out of the 12 combinations of priors and sample periods considered.

Indeed, a comparison of the R^2_{OS} and LPS obtained with the noninformative prior and the more informative prior on ρ_{uw} for each system suggests that ρ_{uw} is more likely to be negative. For instance, in the longest out-of-sample period 1975-2012 (first panel of Table

[20] Gneiting (2011) showed that the MSE is a consistent performance measure when the point prediction equals the mean of the predictive distribution.

1.4), the results obtained with the less predictability prior (prior on R^2 leading to the best results for both systems on this period), for the AR system, using the more informative prior on ρ_{uw} instead of the noninformative, leads to an increase of the R^2_{OS} from -1.03% to 0.28%. Although the sign of the latter R^2_{OS} is positive, the AR system predictions do not outperform significantly the prevailing historical average in terms of MSE. Regarding the CIR system, the noninformative prior on ρ_{uw} yields an R^2_{OS} of -0.30%, less negative than the AR system in this configuration but implying prediction errors still greater than the historical mean. On the other hand, using the more informative prior leads to a positive R^2_{OS} of 1.21%, significant at the 5% level. Consequently using the CIR system with the more informative prior on ρ_{uw} produces significantly more accurate predictions than the historical mean of returns, which is not the case when using the AR system in this case. The conclusions regarding the benefits of using the more informative prior on ρ_{uw} are confirmed by the results obtained for each out-of-sample sub period considered, and for each prior on R^2. Regarding Table 1.5, the conclusions are very similar. The LPS obtained with the more informative prior on ρ_{uw} are always greater than the LPS obtained with noninformative prior. In addition, for all out-of-sample periods, considering the more informative prior, the LPS of the CIR system is higher than the LPS of the AR system.

Table 1.6 presents the corresponding out-of-sample prediction error metrics from the forecasts of a one period ahead predictive regression using the dividend yield as a predictor. For the regression, the R^2_{OS} are negative for all periods. Furthermore, in every period considered the LPS predictability performance metric of the two predictive systems is better than for the regression in all configurations except for the AR system on the period 1975-2012.

We also observe on Tables 1.4 and 1.5 that, the highest R^2_{OS} and LPS for the CIR system is systematically greater than those of the AR system (for all periods), confirming results of section 1.3.2. This suggests there exist benefits in terms of out-of-sample prediction in using the CIR system instead of the AR system and the regression on the dividend yield. Hence, the new features of the CIR system are consistent with the dynamic of the unobservable expected return process in our sample.

Moreover, for both systems, we note that the prior on the degree of return predictability (R^2) leading to the highest R^2_{OS} and LPS is different for most out-of-sample periods considered. In other words, assuming that return predictability is lower or higher implies better out-of-sample estimates, depending on the period. This result suggests that the predictability of returns is in fact time-varying, which is in line with the CIR system and with findings in Rapach et al. (2010) and Henkel et al. (2011). This result also suggest that the outperformance of the CIR system with respect to the AR system in our Bayesian analysis may be explained by the fact that the CIR system incorporates expected returns heteroscedasticity.

Additionally, for all out-of-sample periods considered, using the more informative prior on ρ_{uw} leads to positive R^2_{OS} for both models for at least two of the priors on R^2. However, the significant R^2_{OS} are obtained for all periods for the CIR system (with different

predictability priors on R^2), but only for two out-of-sample periods for the AR system (1975-2000 and 1980-2005).

To illustrate the dynamics of expected returns produced by the different systems, the out-of-sample forecasts over the longest period are presented in Figure 1.8. The priors for the predictive systems correspond to more predictability for R^2 and more informative for ρ_{uw}. The range of fluctuations of expected returns are very different between the predictive systems and the regression. For the latter, the forecasts range from -1.47% to 6.14% per quarter. For the predictive systems, the estimates are much more concentrated and more plausible. Expected returns lie between 1.89% and 3.95%, and between 0.38% and 2.88% for the CIR and AR respectively. The CIR estimates appear to exhibit larger variations during economic recessions than the AR, as suggested by the model. Interestingly, in Figure 1.8 we also observe that both the CIR system and the dividend yield regression produce increasing forecasts at the end of the recession period 2000-2012, while the AR system generates decreasing estimate values. Note that the considerable divergence between the two predictive systems at the beginning of the period can be explained in particular by the difference in posterior mean for the parameter E_r at this first date: 2.80% for the CIR versus 1.30% for the AR. We discuss below the disparities between parameter estimates between the two systems. In addition to the structural differences implied by the models, these parameter divergences contribute to the differences between the expected return time series over the sample.

Figure 1.9 presents the evolution of the posterior mean of the parameters β and ρ_{uw} for the longest out-of-sample period 1975-2012 (re-estimated on the first available date of each year in the sample) using the less predictability prior[21] on R^2 and the more informative on ρ_{uw}. We observe first that the average levels of these parameters differ for the AR and the CIR systems, and second that the posterior means are stable throughout the out-of-sample period for the AR system whereas they vary over time for the CIR system. We notice on Panels A and B of Figure 1.9 that the posterior mean of β for the CIR system (ranging from about 0.8 to 0.4) is always lower than its equivalent in the AR system (which is stable around 0.9). This observation is in line with Kelly and Pruitt (2013)'s findings, and suggests a less persistent expected return process for the CIR system, confirming the results obtained using point estimate parameters in section 1.3.2. Moreover, unlike for the AR system, the posterior mean of β for the CIR system is obviously not constant over time: β is stable around 0.6 from 1975 to 1982, increases to 0.7 by 1986, and then declines smoothly to 0.4 by 2012. This decrease in the level of β explains the progressively lower R^2_{OS} observed in the latest two sub periods (1980-2005 and 1985-2010). To see this, notice that according to equation (1.15), the lower β, the lower the proportion of returns variance explained by expected returns.

Furthermore, Figure 1.10 presents the evolution of the sample autocorrelation of returns at lag 1, using the available returns sample at each point in time. The similarity with the Panel B of Figure 1.9 is striking, as both the sample autocorrelation and the estimated value of β show very similar variation, particularly visible in the recent decline on their

[21] Conclusions are similar using the benchmark prior and more predictability priors on R^2 (unreported).

value. This suggests that the estimation of CIR system captures the fact that returns autocorrelation is not constant over time through the parameter β. Indeed, as observed in equation (1.17) above, there exists an analytic relation between the auto-correlation of returns and the persistence parameter β. A decrease in the latter is related to a decrease in the persistence parameter of expected returns, which is empirically observed only in our estimation of the CIR system.

Another divergence between the systems is observed on Panels C and D of Figure 1.9, where the posterior mean of ρ_{uw} is more negative for the AR system: between -0.6 and -0.8, than for the CIR system: between -0.2 and -0.5. This observation also confirms the results obtained in section 1.3.2: the value of ρ_{uw} should be closer to zero if we adopt the CIR system as the predictive model. Interestingly, similarly to the parameter β, the posterior mean of ρ_{uw} is relatively constant over time for the AR system while it is time-varying for the CIR system. Indeed, for the latter, the variations of ρ_{uw} in Panel D are virtually the mirror image of the variations of β in Panel B (sample correlation of -0.86 between the posteriors of β and ρ_{uw}), meaning that empirically, as β decreases, the correlation between expected and unexpected returns becomes less negative.

These conclusions are confirmed by Figure 1.11, which presents the posterior distributions of parameters ρ_{uw} and β, corresponding to the estimations for 1992 and twenty years later in 2012 (last estimation). While the posterior distributions for the AR system are substantially the same, there exist clear differences for the CIR system between the two dates of estimation. Posteriors for ρ_{uw} are shifted to zero (less negative) between 1992 and 2012, as well as posteriors for β (much less close to one). Indeed, for the CIR system, in 1992, most of the mass of the distribution is below zero for ρ_{uw}, and between 0.6 and 1 for β. This is no longer the case in 2012, when the distribution of ρ_{uw} is much more neutral (less clearly negative, although the negative part is still heavier than the positive one), and posterior values of β are spread from 0 to 1 with a mode around 0.5.

In summary, the evolution of the posterior distributions and posterior means of the CIR system parameters highlight two differences with the AR system. First, the shape of the posterior distribution of β actually changes over the period studied, and the persistence of expected return steadily decreases during the last thirty years of our sample. This behavior is consistent with the autocorrelation of realized returns (at lag 1) which presents a similar decreasing trend over the period, as expected given equation (1.17). Second, the posterior distribution of ρ_{uw} also varies throughout the sample, and it gradually becomes less negative toward the end of the sample.

Conclusion

A widely held prior belief about the unobservable expected return of the stock market is that it should have similar features to economically motivated predictors such as the dividend yield, thus being a stable quantity moving around its long-term mean in a persistent and counter-cyclical way. Consequently, expected returns are often modeled as an autoregressive process of order one with counter-cyclical variations with respect to market

returns.

Empirical studies using economic predictors such as Rapach et al. (2010), and Henkel et al. (2011) have found that return predictability in the stock market is stronger during recessions. The latter finding implies that the conditional variance of expected returns must increase in economic downturns. This fact contrasts with the constant conditional variance of the standard autoregressive process (AR) of order one often used to model expected returns.

In this paper we explore the implications of modeling expected returns with a discretized CIR process within a predictive system framework. The CIR model induces a continuously changing conditional variance for expected returns, increasing during market downturns, which implies a countercyclical degree of return predictability. Indeed, the CIR system without predictors reproduces the counter-cyclical variations in the predictive power of the dividend-price ratio documented by Rapach et al. (2010), and Henkel et al. (2011), without any additional parameters relative to the AR model. The CIR model also implies a natural positive relationship between expected returns and the expected variance of realized returns, i.e., a positive risk-return tradeoff, which is absent in the AR model.

Additional theoretical and empirical implications of the CIR system are that the expected returns are unlikely to be negative (consistent with the results of empirical studies such as Campbell and Thompson, 2008; Pettenuzzo et al., 2014), and that the conditional covariance between expected and realized returns is time-varying (higher in absolute value during bad times). In addition, the CIR system implies a time-varying conditional autocovariance of realized returns that is higher during recessions (when expected returns are high), and lower during expansions (when expected returns are low). Furthermore, in out-of-sample tests we find that the CIR system without any predictor can produce significantly better return predictions than the historical average, and improvements in forecasts relative to both the predictive regression using the dividend yield, and the AR predictive system without predictors.

In our Bayesian parameter estimation of the CIR system, we also find that the persistence of expected returns is lower than suggested by previous studies, and is not constant over time. It has steadily decreased for the last two decades, in tandem with the sample autocorrelation of realized returns.

Tables and Figures

R^2	E_r	σ_r	β	k-e ρ_{uw}	σ_μ	$\sigma_{w,ar}$	$\sigma_{w,cir}$	σ_u
0.09%	1.77%	8.38%	0.90	-0.06	0.25%	0.11%	0.82%	8.38%
5%	1.77%	8.38%	0.90	-0.47	1.87%	0.82%	6.14%	8.17%

Table 1.1: Point estimate parameters for different levels of the R^2 of the regression of r_{t+1} on μ_t. E_r and σ_r correspond to mean and standard deviation sample estimates for quarterly returns of the CRSP aggregate US market index from 1952 to 2012, and β is taken as in Pástor and Stambaugh (2009) section II.B. The following columns follow from the respective definitions in the text which are functions of the first four columns; k-e ρ_{uw} stands for knife-edge value of ρ_{uw}. Given a value for R^2 and assuming $\mathbb{E}(\sqrt{\mu_{t-1}}) \approx \sqrt{E_r}$, the knife-hedge value of ρ_{uw} is the same for both systems.

$\boldsymbol{\rho_{uw} = -0.47}$	Intercept	Slope	R^2
CIR system	0.0177	-	0.00
AR system	0.0177	-	0.00
$\boldsymbol{\rho_{uw} = -0.66}$	Intercept	Slope	R^2
CIR system	-0.597	7.35	0.470
	(-14.2)	(14.6)	
AR system	1.65	-19.5	0.00312
	(0.879)	(-0.870)	
$\boldsymbol{\rho_{uw} = -0.85}$	Intercept	Slope	R^2
CIR system	-0.598	7.39	0.460
	(-13.9)	(14.3)	
AR system	1.37	-16.2	0.0120
	(1.73)	(-1.71)	

Table 1.2: Results of the regression of expected returns on expected volatility of returns. This table summarizes the results from regressions: $\mathbb{E}(r_{t+1}|D_t) = a + b\sqrt{\text{Var}(r_{t+1}|D_t)} + e_t$, where $\mathbb{E}(r_{t+1}|D_t) = \mathbb{E}(\mu_t|D_t)$ is the filtered expected return from the predictive systems, and $\text{Var}(r_{t+1}|D_t)$ is the expected variance of returns estimated by the systems. The table reports the intercept, the slope of the regression and the (unadjusted) R^2 for several values of ρ_{uw} (the correlation between expected and unexpected returns). t-statistics are given in parentheses. The case $\rho_{uw} = -0.47$ corresponds to historical average as estimate of expected returns. For the Kalman filtering procedure, the (unconditional) variance of μ is set to 5% of the variance of r.

1975-2012	R^2_{OS}	β	R^2	ρ_{uw}
		AR system		
Best configuration	3.00%	0.9	4%	-0.95
Worst configuration	-5.81%	0.9	5%	0.95
		CIR system		
Best configuration	3.39%	0.8	2%	-0.95
Worst configuration	-5.62%	0.8	5%	0.95

Table 1.3: Out-of-sample results summary with point estimate parameters. Each line presents the highest (or lowest) R^2_{OS} obtained for each predictive system (AR or CIR), as well as the corresponding parameter set. Predictions are computed on quarterly returns on the value-weighted portfolio of all NYSE, Amex, and Nasdaq stocks in excess of the quarterly return on a 1-month T-bill obtained from CRSP. The sample begins in 1952 and the out-of-sample period is 1975-2012. We use the prevailing returns average for E_r.

1975-2012	**Less predictability**		**Benchmark prior**		**More predictability**	
Prior on ρ_{uw}	Noninf.	More inf.	Noninf.	More inf.	Noninf.	More inf.
CIR R^2_{OS} (%)	-0.30	1.21**	-0.73	1.07**	-1.26	0.71*
AR R^2_{OS} (%)	-1.03	0.28	-1.54	0.13	-2.23	-0.23
1975-2000	**Less predictability**		**Benchmark prior**		**More predictability**	
Prior on ρ_{uw}	Noninf.	More inf.	Noninf.	More inf.	Noninf.	More inf.
CIR R^2_{OS} (%)	-0.02	2.64***	-0.29	2.93***	-0.66	3.01***
AR R^2_{OS} (%)	-1.41	0.69*	-2.14	0.61*	-3.13	0.18
1980-2005	**Less predictability**		**Benchmark prior**		**More predictability**	
Prior on ρ_{uw}	Noninf.	More inf.	Noninf.	More inf.	Noninf.	More inf.
CIR R^2_{OS} (%)	0.41	1.48**	0.11	1.51**	-0.29	1.34**
AR R^2_{OS} (%)	-0.12	0.52	-0.34	0.62*	-0.72	0.66*
1985-2010	**Less predictability**		**Benchmark prior**		**More predictability**	
Prior on ρ_{uw}	Noninf.	More inf.	Noninf.	More inf.	Noninf.	More inf.
CIR R^2_{OS} (%)	0.04	0.80*	-0.41	0.67	-0.96	0.30
AR R^2_{OS} (%)	-0.20	0.04	-0.39	0.04	-0.73	0.02

Table 1.4: Out-of-sample results of the predictive systems using the Bayesian procedure. For each out-of-sample period, for both predictive systems, three priors on the R^2 of the regression of r_{t+1} on μ_t (less predictability, benchmark prior and more predictability) and two priors on ρ_{uw} (noninformative and more informative) are applied. Predictions are computed on quarterly returns on the value-weighted portfolio of all NYSE, Amex, and Nasdaq stocks in excess of the quarterly return on a 1-month T-bill obtained from CRSP. Our sample begins in 1952-Q1. The predictive systems parameters are re-estimated on the first available date of each year in the sample. Predictions are computed each new quarter using the data available at each point in time.

1975-2012	**Less predictability**		**Benchmark prior**		**More predictability**	
Prior on ρ_{uw}	Noninf.	More inf.	Noninf.	More inf.	Noninf.	More inf.
CIR LPS	1.01	1.019	1.007	1.018	1.004	1.015
AR LPS	1.006	1.014	1.002	1.012	0.997	1.007
1975-2000	**Less predictability**		**Benchmark prior**		**More predictability**	
Prior on ρ_{uw}	Noninf.	More inf.	Noninf.	More inf.	Noninf.	More inf.
CIR LPS	1.062	1.076	1.060	1.077	1.058	1.078
AR LPS	1.054	1.065	1.049	1.063	1.043	1.059
1980-2005	**Less predictability**		**Benchmark prior**		**More predictability**	
Prior on ρ_{uw}	Noninf.	More inf.	Noninf.	More inf.	Noninf.	More inf.
CIR LPS	1.041	1.046	1.039	1.046	1.036	1.045
AR LPS	1.038	1.041	1.036	1.041	1.033	1.040
1985-2010	**Less predictability**		**Benchmark prior**		**More predictability**	
Prior on ρ_{uw}	Noninf.	More inf.	Noninf.	More inf.	Noninf.	More inf.
CIR LPS	1.007	1.012	1.004	1.011	1.001	1.008
AR LPS	1.007	1.008	1.005	1.007	1.002	1.006

Table 1.5: Log predictive scores (LPR) of the predictive systems using the Bayesian procedure. For each out-of-sample period, for both predictive systems, three priors on the R^2 of the regression of r_{t+1} on μ_t (less predictability, benchmark prior and more predictability) and two priors on ρ_{uw} (noninformative and more informative) are applied. Predictions are computed on quarterly returns on the value-weighted portfolio of all NYSE, Amex, and Nasdaq stocks in excess of the quarterly return on a 1-month T-bill obtained from CRSP. Our sample begins in 1952-Q1. The predictive systems parameters are re-estimated on the first available date of each year in the sample. Predictions are computed each new quarter using the data available at each point in time.

Div. Yield	**1975-2012**	**1975-2000**	**1980-2005**	**1985-2010**
R^2_{OS} (%)	-0.012	-0.036	-0.045	-0.041
LPS	1.003	1.04	1.011	0.982

Table 1.6: Out-of-sample results of the predictive regression using the dividend yield. Predictions are computed on quarterly returns on the value-weighted portfolio of all NYSE, Amex, and Nasdaq stocks in excess of the quarterly return on a 1-month T-bill obtained from CRSP. Our sample begins in 1952-Q1. We use a growing window. The predictive regression parameters are re-estimated each new quarter and predictions are computed using the data available at each point in time.

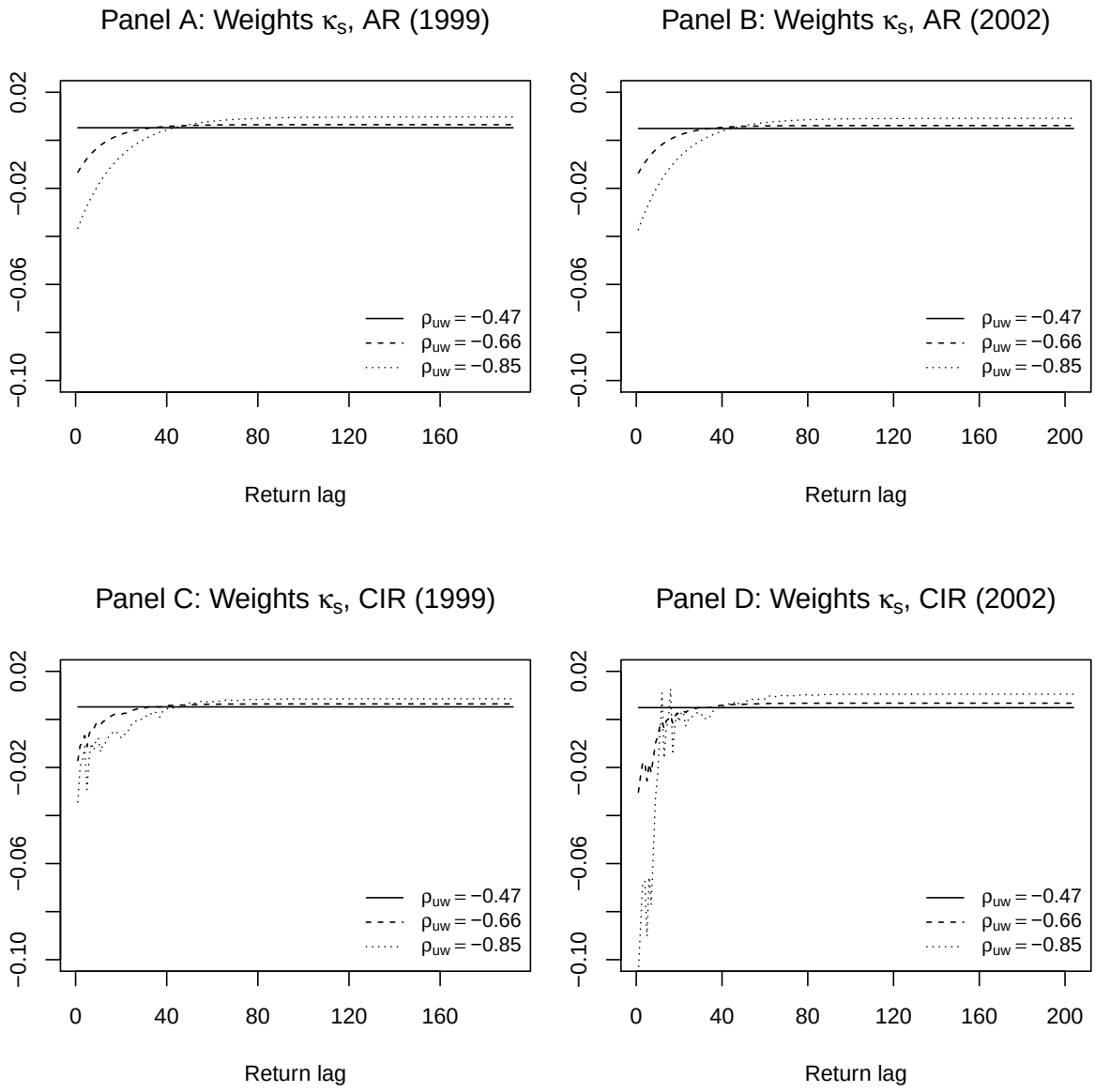

Figure 1.1: The effect of lagged returns on $E(r_{t+1}|D_t)$ when no predictors are used. This figure plots finite-sample values of κ_s, the weights on lagged total returns in $E(r_{t+1}|D_t)$. The samples considered are 1952Q1-1999Q4 for Panels A and C, and 1952Q1-2002Q4 for Panels B and D. The autoregressive coefficient is set to $\beta = 0.9$ and the unconditional mean return E_r is estimated by the sample mean over the quarters since 1952Q1. In both Panels, the predictive R^2 corresponding to the fraction of variation in r_{t+1} that can be explained by μ_t is set to 0.05.

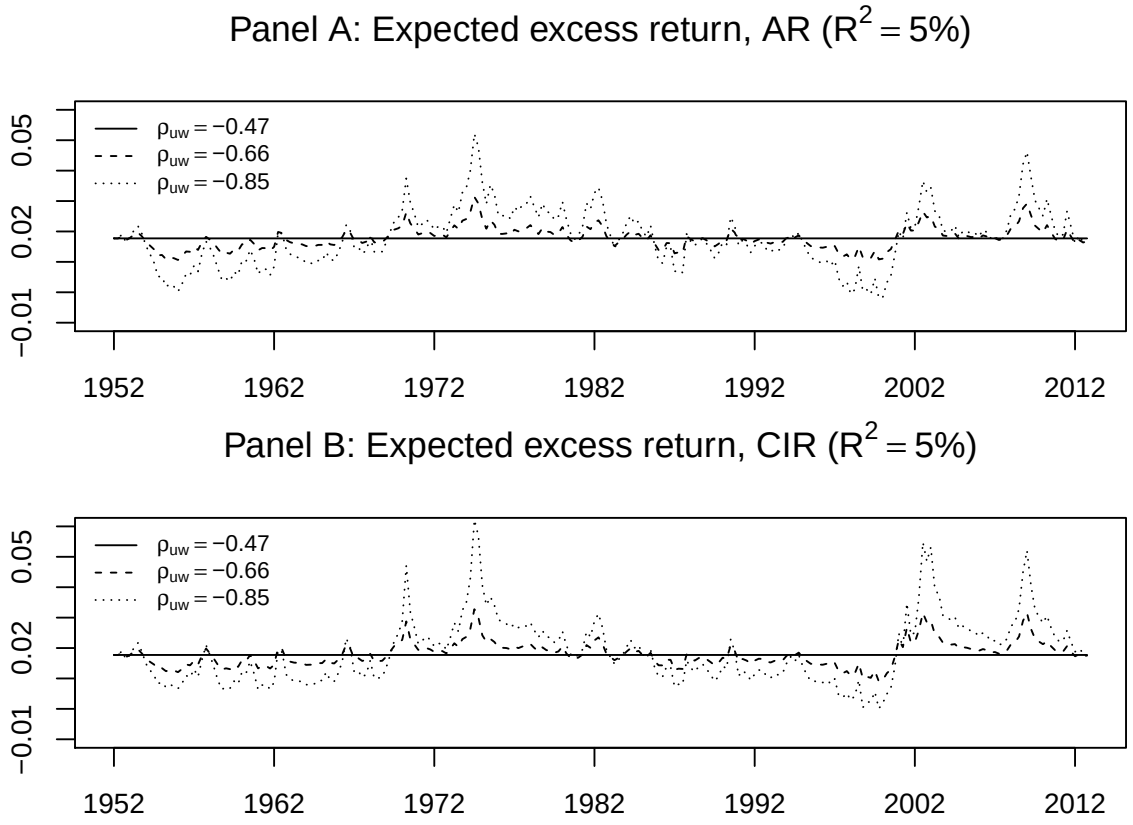

Figure 1.2: Panel A presents the equity premium $\mathbb{E}(r_{t+1}|D_t) = \mathbb{E}(\mu_t|D_t)$ from Pastor and Stambaugh's AR(1) predictive system, and Panel B the expected excess returns from our CIR-type predictive system. This figure displays the time series of quarterly US stock market premium from 1952Q1 to 2012Q4 estimated for different values of ρ_{uw} (the flat line corresponds to the knife-hedge value of ρ_{uw}, i.e., historical average as estimate of expected return). The autoregressive coefficient is set to $\beta = 0.9$ and the unconditional mean return E_r is estimated by the sample mean over the whole sample. In both panels, the R^2 corresponding to the fraction of variation in r_{t+1} that can be explained by μ_t is set to 5%.

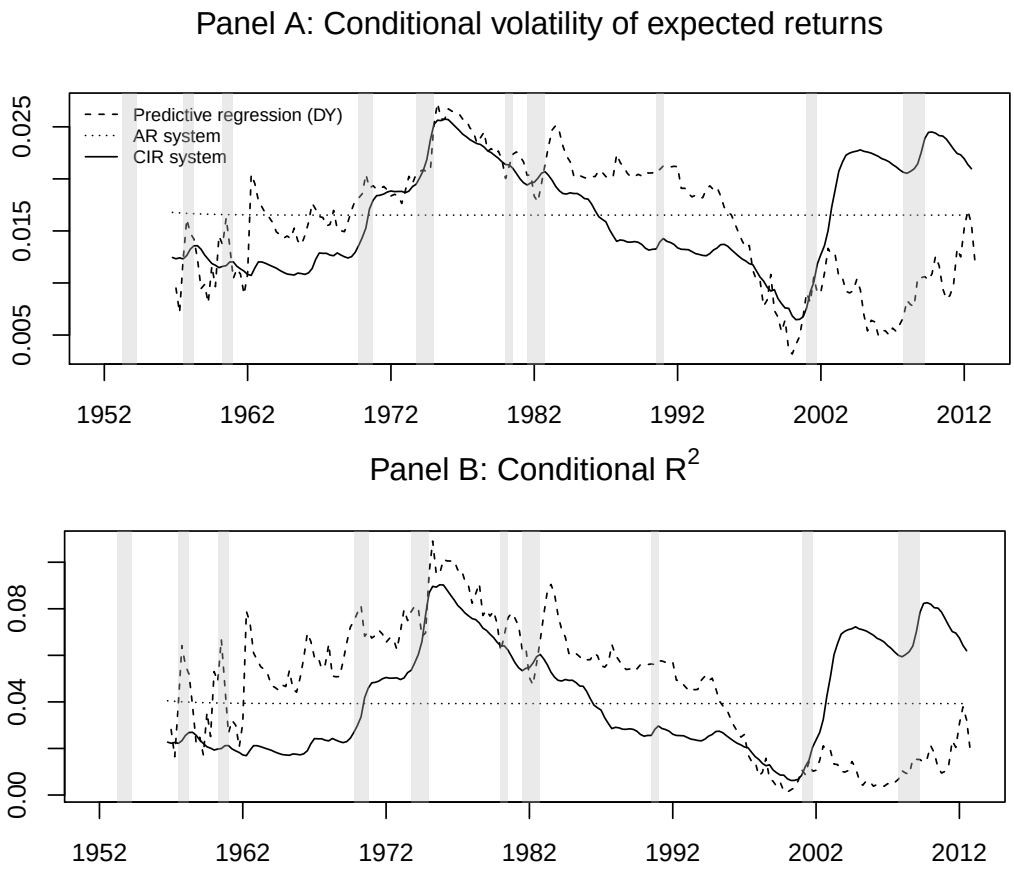

Figure 1.3: Panel A plots the square root of estimates of $\text{Var}(\mu_t|D_t)$, where μ_t denotes the expected stock return from time t to time $t+1$ and D_t denotes the information set observed through time t. The conditional variance of μ is presented for the CIR system and the AR system without predictors and for the predictive regression using the dividend price-ratio as predictor. The sample considered is 1952Q1-2012Q4. The parameters used for the predictive systems corresponds to $R^2 = 5\%$, $\rho_{uw} = -0.85$ and $\beta = 0.9$. The unconditional mean E_r is estimated with the sample mean over the whole period. Panel B presents the ratio of the conditional variances of μ_t and r_{t+1} for both systems and the predictive regression. The grayed areas correspond to economic recessions as reported by NBER.

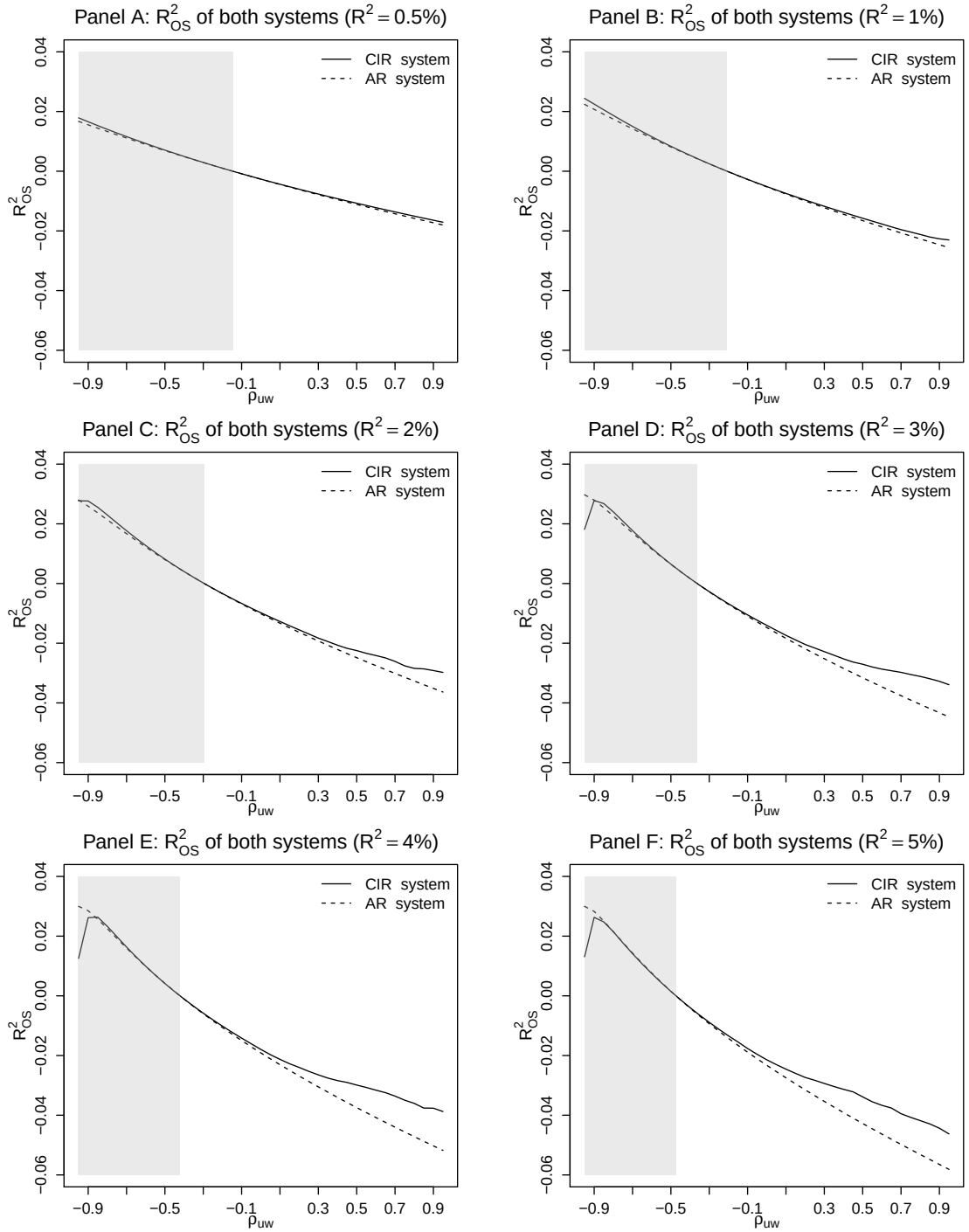

Figure 1.4: Out-of-sample results with point estimate parameters. Each Panel presents the R^2_{OS} as a function of ρ_{uw} when the R^2 of the regression of r_{t+1} on μ_t is set to a given value. Predictions are computed on quarterly returns on the value-weighted portfolio of all NYSE, Amex, and Nasdaq stocks in excess of the quarterly return on a 1-month T-bill obtained from CRSP. The sample begins in 1952 and the out-of-sample period is 1975-2012. We use a constant β of 0.9 and the prevailing returns average for E_r. The grayed areas correspond to values of ρ_{uw} implying a dominant change effect (countercyclical expected returns).

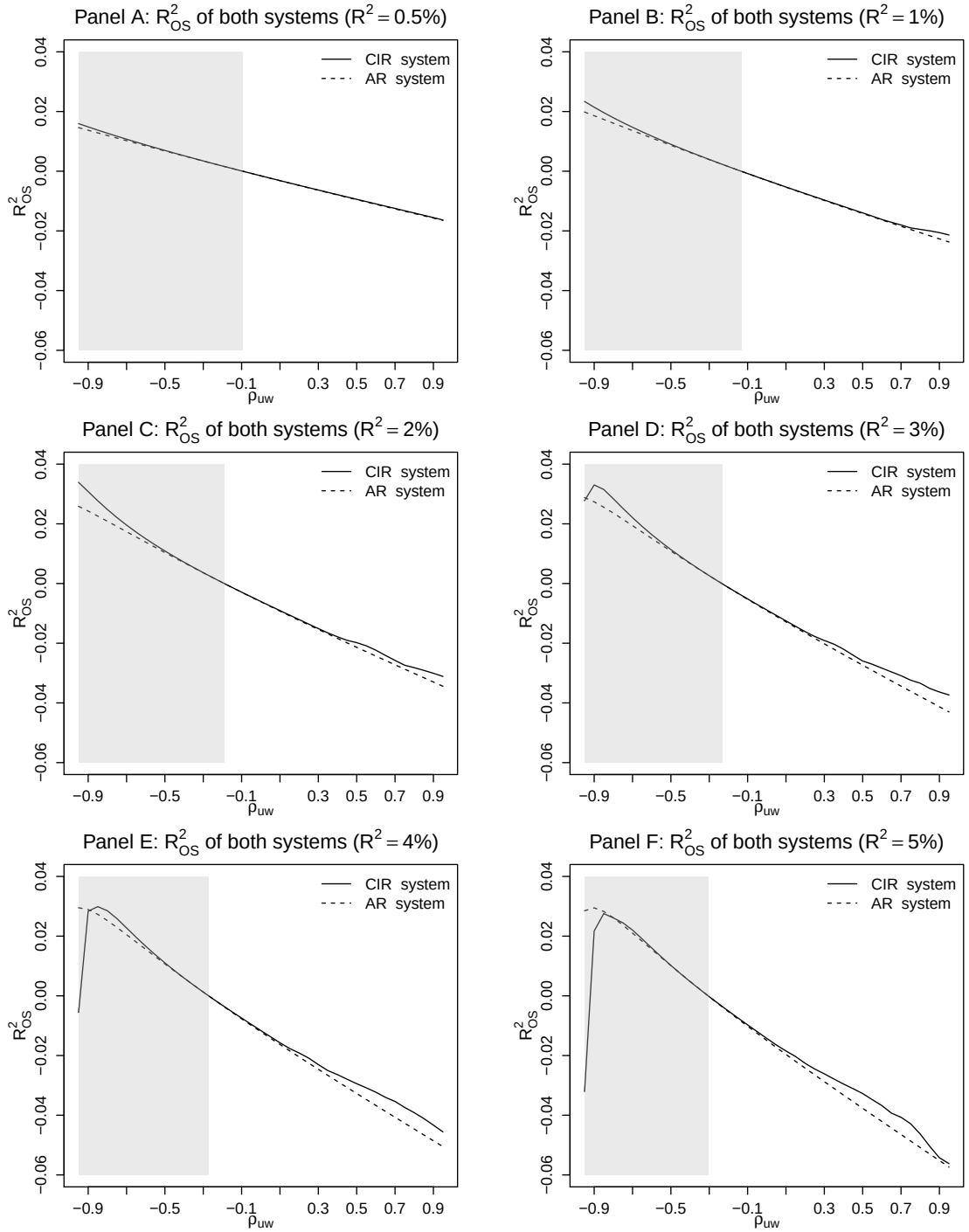

Figure 1.5: Out-of-sample results with point estimate parameters. Each Panel presents the R^2_{OS} as a function of ρ_{uw} when the R^2 of the regression of r_{t+1} on μ_t is set to a given value. Predictions are computed on quarterly returns on the value-weighted portfolio of all NYSE, Amex, and Nasdaq stocks in excess of the quarterly return on a 1-month T-bill obtained from CRSP. The sample begins in 1952 and the out-of-sample period is 1975-2012. We use a constant β of 0.8 and the prevailing returns average for E_r. The grayed areas correspond to values of ρ_{uw} implying a dominant change effect (countercyclical expected returns).

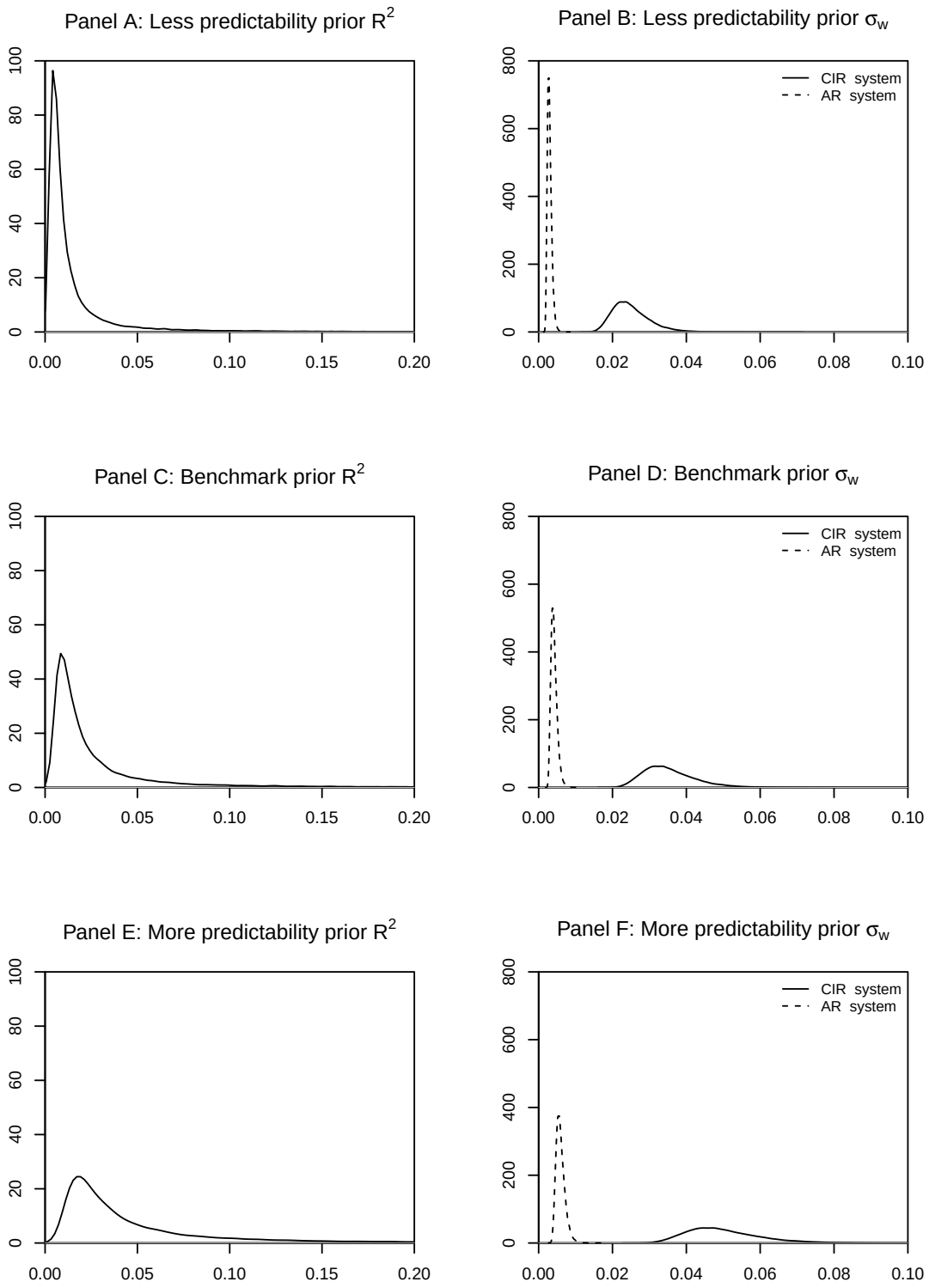

Figure 1.6: The prior distributions for R^2 and σ_w used in the Bayesian analysis. Panels A and B plot the prior on the R^2 from the regression of r_{t+1} on μ_t and the corresponding prior on σ_w for both AR system and CIR system, corresponding to less predictability prior. Panels C and D plot the distributions corresponding to the benchmark prior. Panels E and F plot the more predictability prior distributions.

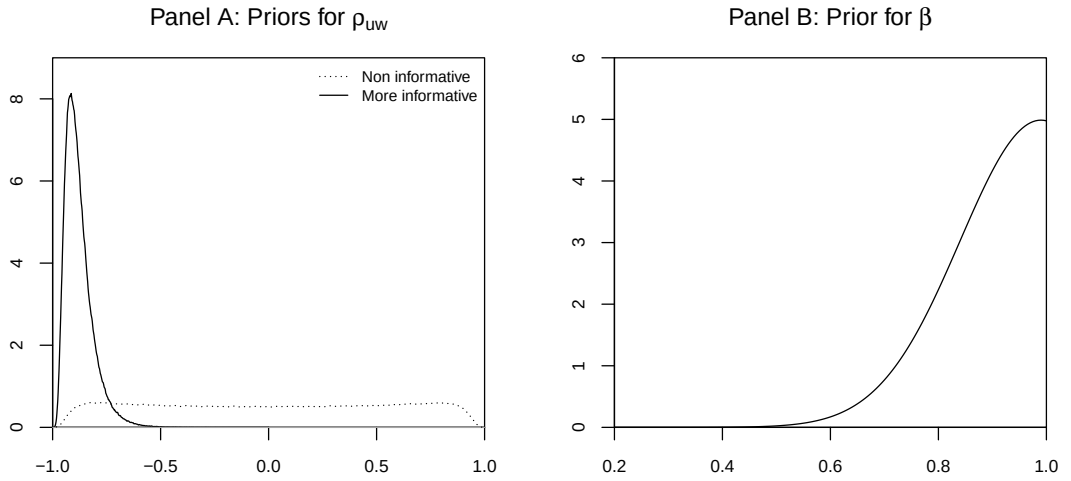

Figure 1.7: The prior distributions for ρ_{uw}, β and the implied priors of ρ_{uw}^2 used in the Bayesian analysis. Panel A plots the two prior distributions for ρ_{uw}: noninformative (flat between -0.9 and 0.9) and more informative (most of the mass below -0.71). Panel B plots the prior on the autoregressive coefficient β in the dynamics of μ.

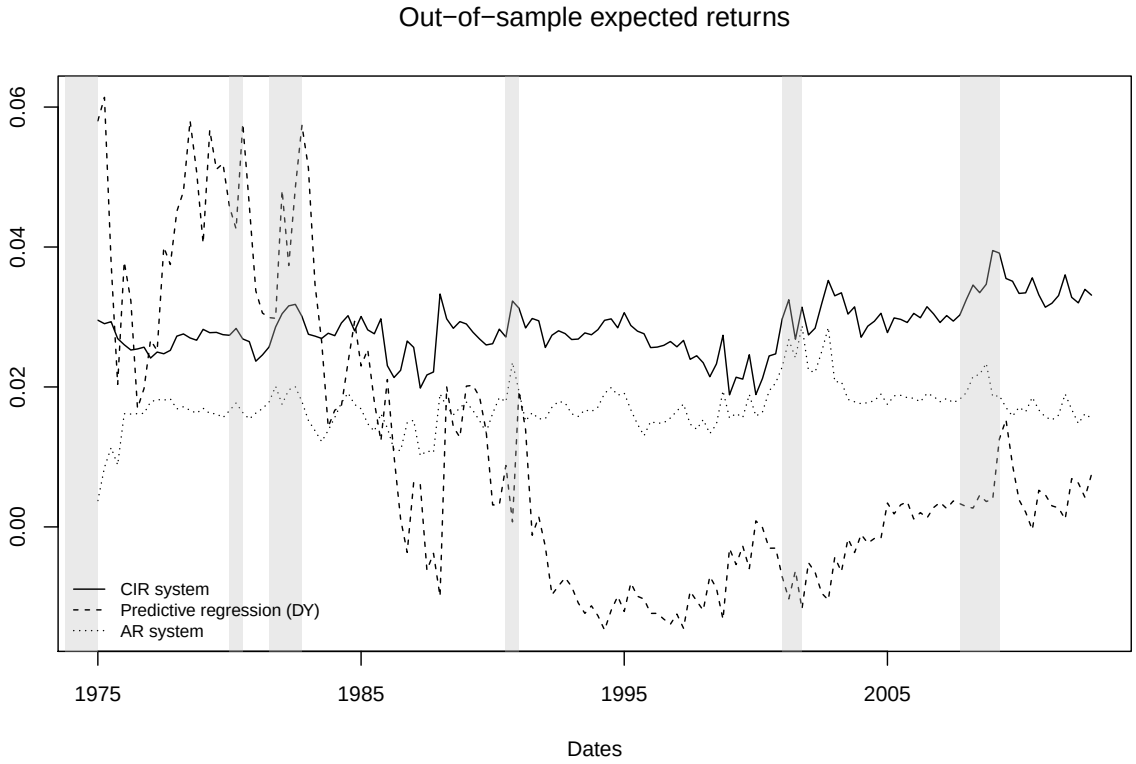

Figure 1.8: The expected excess returns. The figure plots the out-of-sample quarterly expected returns estimated by the CIR predictive system, the OLS predictive regression using the dividend yield, and the AR system. Predictions are computed on quarterly returns on the value-weighted portfolio of all NYSE, Amex, and Nasdaq stocks in excess of the quarterly return on a 1-month T-bill obtained from CRSP. The sample begins in 1952 and the out-of-sample period is 1975-2012. The grayed areas correspond to economic recessions as reported by NBER. Priors for the predictive systems correspond to more predictability for R^2 and more informative for ρ_{uw}.

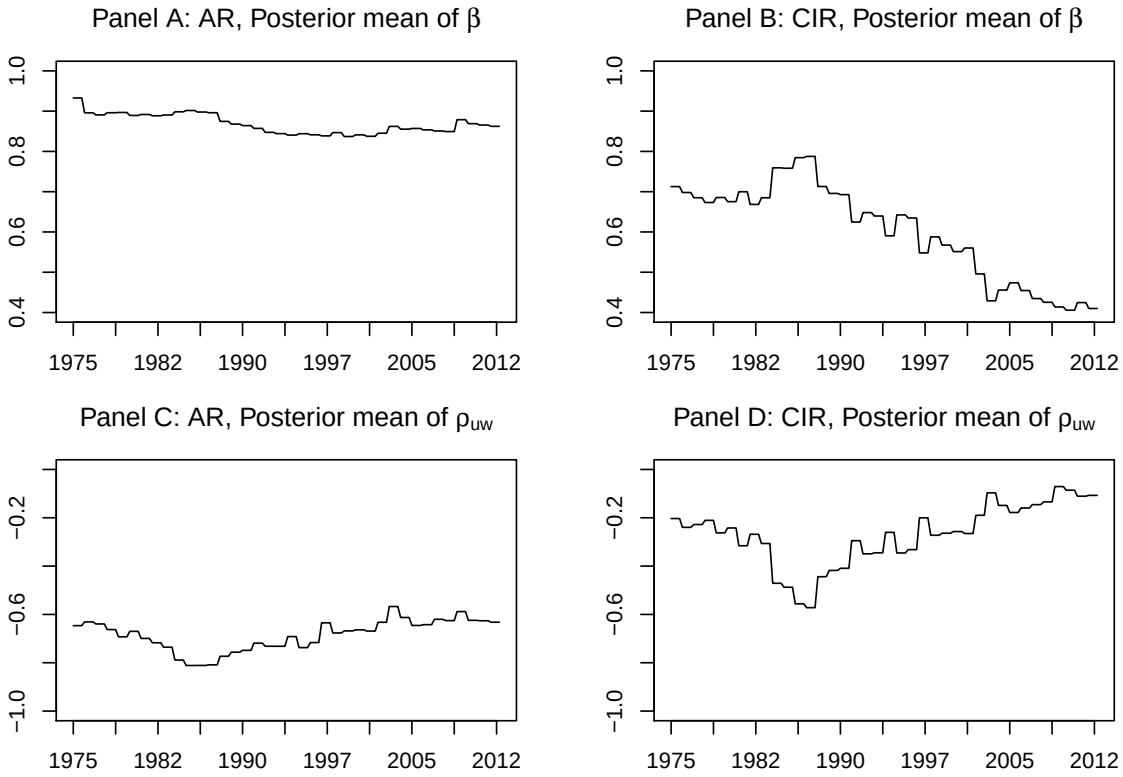

Figure 1.9: Evolution of the mean of the posterior distributions of β, ρ_{uw}, and ρ_{uw}^2 for the AR and CIR predictive systems. The priors are more informative on ρ_{uw} and correspond to less predictability prior on R^2. The considered out-of-sample period is 1975Q1-2012Q4. Predictive systems are re-estimated on the first available date of each year in the sample.

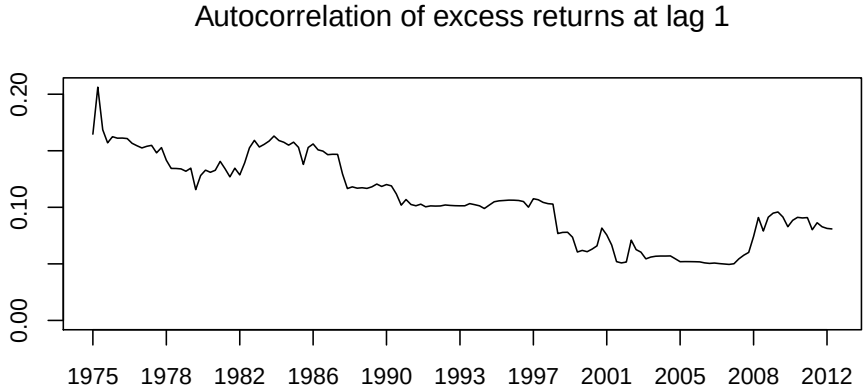

Figure 1.10: Evolution of the autocorrelation at lag 1 of returns. Our sample begins in 1952 and corresponds to quarterly returns on the value-weighted portfolio of all NYSE, Amex, and Nasdaq stocks in excess of the quarterly return on a 1-month T-bill obtained from CRSP. Each quarter, we re-estimate the autocorrelation at lag 1 using the available returns.

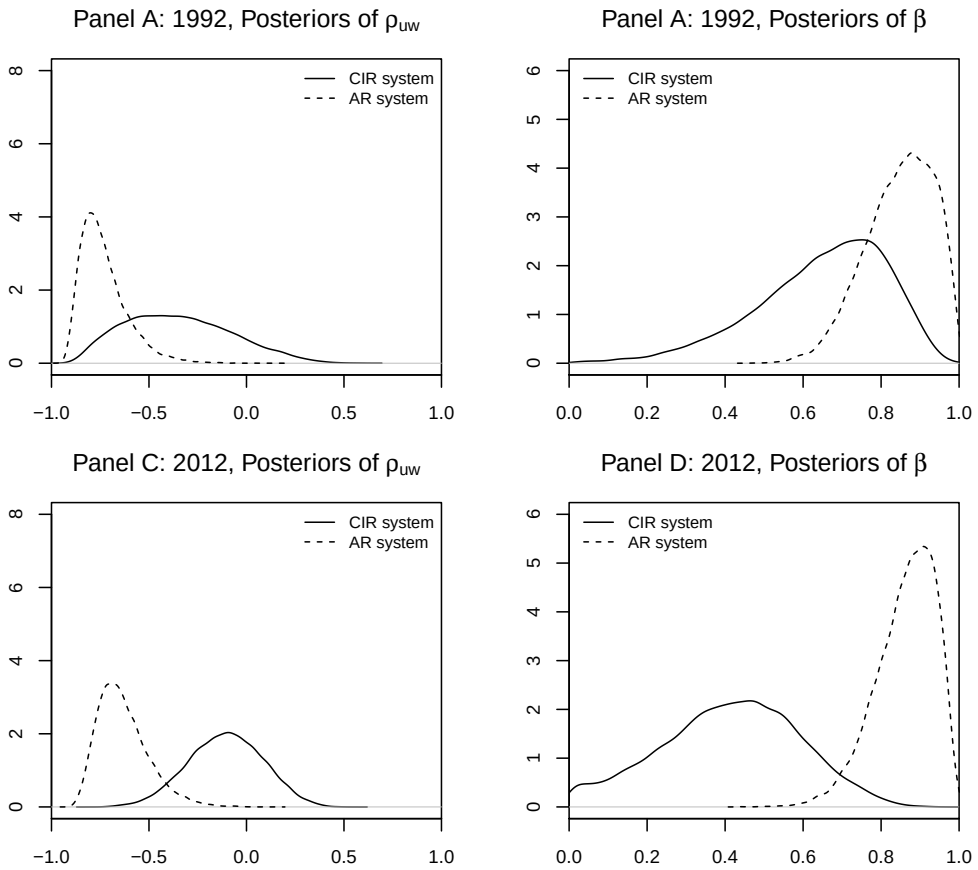

Figure 1.11: Posterior distributions of β and ρ_{uw} for the AR and CIR predictive systems. The priors are more informative on ρ_{uw} and correspond to less predictability prior on R^2. The posterior distributions are obtained from systems estimation in 1992-Q1 and 2012-Q1.

Appendix

1.A Discretization of CIR process

The Cox et al. (1985) (CIR) model is defined by the following Stochastic Differential Equation (SDE):

$$dX_t = \kappa(\theta - X_t)dt + \sigma\sqrt{X_t}dW_t, \quad X_0 \geq 0, \tag{1.39}$$

where κ, θ and σ are constants, and W is a standard Brownian motion. This SDE has a level dependent diffusion term ($\sigma\sqrt{X_t}$) implying a conditional heteroscedasticity for the process X. Furthermore, the model can rule out negative values for X. Indeed, given that κ is a Lipschitz constant for the drift term of (1.39), Feller's test (c.f. Feller, 1951) for univariate stochastic process ensures that the following condition $\kappa\theta \geq \frac{\sigma^2}{2}$ implies $\mathbb{P}(\tau_0^x = \infty) = 1$, where $\tau_0^x = \inf\{t \geq 0 : X_t = 0\}$ and x refers to the case $X_0 = x \geq 0$. Details are given in Berkaoui, Bossy, Diop, et al. (2008).

Although the exact non-central chi-squared distribution of the increments of the CIR equation (1.39) could be used in a discrete time framework (as in Gouriéroux and Jasiak, 2006), the study of the correlation structure between this distribution and the Gaussian innovations of realized returns would require a non-Gaussian copula and/or numerical methods, which is inconvenient to develop the economic intuition of the model. A direct Euler discretization with time step Δt of (1.39) is:

$$\mu_{t+\Delta t} = \mu_t + \kappa(\theta - \mu_t)\Delta t + \sigma\sqrt{\mu_t}(W_{t+\Delta t} - W_t). \tag{1.40}$$

However process (1.40) does not have a *strictly* zero probability of being negative as the Gaussian increment is not bounded from below and it is important to notice that the term $\sqrt{\mu_t}$ is not defined for a given $\mu_t < 0$. Hence in our numerical filtering procedures (application of algorithm described in section 1.C), we use instead the following discretization:

$$\mu_{t+\Delta t} = \mu_t + \kappa(\theta - \mu_t)\Delta t + \sigma\sqrt{|\mu_t|}(W_{t+\Delta t} - W_t). \tag{1.41}$$

Berkaoui et al. (2008) show that for the process defined by (1.41), for all t given $\mu_{\eta(t)} = x$, where x is a positive value and $\eta(t)$ is the closest previous step of discretization, the following probability inequality applies

$$\mathbb{P}^x_{cir} = \mathbb{P}(\mu_t \leq 0 \mid \mu_{\eta(t)} = x) \leq \frac{1}{2}\exp\left(-\frac{x}{8\sigma^2\Delta t}\right); \tag{1.42}$$

thus, x being sufficiently above zero and/or taking time steps Δt sufficiently small, greatly lowers the chance of μ_t of becoming negative. Notice that, for the AR system, the value of the conditional probability $\mathbb{P}(\mu_t \leq 0 \mid \mu_{t-1} = x)$ has the following expression if we assume that μ has the AR(1) dynamics given by equation (1.3) in the text:

$$\mathbb{P}^x_{ar} = \mathbb{P}\left(w_t \leq -(1-\beta)E_r - \beta x\right) = \Phi\left(\frac{-(1-\beta)E_r - \beta x}{\sigma_{w,ar}}\right), \tag{1.43}$$

where Φ stands for the cumulative distribution function of a standard Gaussian random variable.

Throughout our economical analysis we assume that actual realizations of the process μ, given by equation (1.40), are non-negative. This is a reasonable assumption given the fact that μ is an unobservable process and that the non-negativity assumption is one of the motivations for using a discretized CIR process instead of the AR(1) process. On the other hand, if the aim would be to simulate the CIR-type process μ_t (instead of estimating it), a more adapted equation for this purpose would be for instance the symmetrized Euler scheme of (1.39), studied by Bossy, Diop, et al. (2007):

$$\hat{\mu}_{t+\Delta t} = \left| \hat{\mu}_t + \kappa(\theta - \hat{\mu}_t)\Delta t + \sigma\sqrt{\hat{\mu}_t}(W_{t+\Delta t} - W_t) \right|, \tag{1.44}$$

which ensures the positivity of the discretized process[22]. However discretization (1.44) does not allow to explore all the economic implications of the model, as we are unable to derive closed-form solutions for the autocovariance of returns (equation 1.11 in the text) and the variance of μ, for instance. Thus, in order to develop the economic implications of the expected return positivity condition on the predictive system we work with the CIR discretization (1.40) and assume that any actual path of μ remains positive in our analysis.

1.B Autocovariance of returns

In this section we derive the autocovariance of returns using equation (1.1) in the text. We assume first the following (more general) dynamics for μ:

$$\mu_{t+1} = (1-\beta)E_r + \beta\mu_t + g(\mu_t)w_{t+1}, \tag{1.45}$$

where g is a general function including the case $g(.) = \sqrt{(.)}$ or $g(.) = \sqrt{|.|}$.
Notice that the autocovariance of returns at lag $k \geq 0$ is:

$$\begin{aligned} \mathrm{Cov}(r_{t+k}, r_t) &= \mathrm{Cov}(\mu_{t+k-1} + u_{t+k},\ \mu_{t-1} + u_t) \\ &= \mathrm{Cov}(\mu_{t+k-1}, \mu_{t-1}) + \mathrm{Cov}(\mu_{t+k-1}, u_t). \end{aligned} \tag{1.46}$$

The first term on the right hand side of equation (1.46) is the autocovariance of expected returns at lag k. Using equation (1.9) (with $\sqrt{(.)}$ expressed as $g(.)$) the autocovariance

[22] Alfonsi (2005) discusses several discretization schemes for the simulation of the CIR process.

of μ is equal to:

$$\begin{aligned}\mathrm{Cov}(\mu_{t+k},\mu_t) &= \mathrm{Cov}\left(\sum_{j=0}^{\infty}\beta^j w_{t+k-j}g(\mu_{t+k-1-j}),\ \sum_{i=0}^{\infty}\beta^i w_{t-i}g(\mu_{t-1-i})\right)\\ &= \sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\beta^{i+j}\mathrm{Cov}\left(w_{t+k-j}g(\mu_{t+k-1-j}),\ w_{t-i}g(\mu_{t-1-i})\right)\\ &= \sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\beta^{i+j}\mathbb{E}\left[w_{t+k-j}g(\mu_{t+k-1-j})\ w_{t-i}g(\mu_{t-1-i})\right].\end{aligned}$$

To see this notice that the product of expectations is zero in the covariance due to the lag between the noise terms and the terms containing $g(\mu)$ (which are independent). Moreover, the only nonzero terms in the sums are when $j = i + k$. Thus we obtain:

$$\mathrm{Cov}(\mu_{t+k},\mu_t) = \sum_{i=0}^{\infty}\beta^{2i+k}\sigma_w^2\mathbb{E}\left[g(\mu_{t-1-i})^2\right] = \beta^k\sigma_w^2\sum_{i=0}^{\infty}\beta^{2i}\mathbb{E}\left[g(\mu_{t-1-i})^2\right]. \tag{1.47}$$

Equation (1.47) easily simplifies to $\beta^k\left(\frac{\sigma_w^2 E_r}{1-\beta^2}\right)$ and allows us to compute the variance of μ when $g(.) = \sqrt{(.)}$ as in section 1.2.3 of the text. Note that this result is valid if the trajectory of μ stays positive at all time such that the term $\sqrt{\mu_t}$ is well defined for all t. However, as explained above, using a discretization of a CIR process (with a Gaussian noise) introduces a bias that could lead to eventual negative values (with low probability, cf. Appendix 1.A) forcing us to consider the case $g(.) = \sqrt{|.|}$. In order to check that the use of this function $g(.)$ (defined for all real numbers) in our empirical analysis does not alter the analytical properties of the model, we derive here the autocovariance in this particular case which has to be linked with equations (1.11) and (1.12) of the text. The derivation of the second term of the right hand side of (1.46) yields to:

$$\begin{aligned}\mathrm{Cov}(\mu_{t+k-1},u_t) &= \mathrm{Cov}\left(\sum_{i=0}^{\infty}\beta^i w_{t+k-1-i}\sqrt{|\mu_{t+k-2-i}|},\ u_t\right)\\ &= \sum_{i=0}^{\infty}\beta^i\mathrm{Cov}(w_{t+k-1-i}\sqrt{|\mu_{t+k-2-i}|},\ u_t)\\ &= \sum_{i=0}^{\infty}\beta^i\mathbb{E}[w_{t+k-1-i}\sqrt{|\mu_{t+k-2-i}|}\ u_t],\end{aligned}$$

where the only nonzero term is obtained when $i = k - 1$. Thereby,

$$\mathrm{Cov}(\mu_{t+k-1},u_t) = \beta^{k-1}\mathbb{E}[w_t\sqrt{|\mu_{t-1}|}\ u_t] = \beta^{k-1}\sigma_{uw}\mathbb{E}[\sqrt{|\mu_{t-1}|}]. \tag{1.48}$$

The autocovariance of μ is now:

$$\begin{aligned}\mathrm{Cov}(\mu_{t+k},\mu_t) &= \sum_{i=0}^{\infty}\beta^{2i+k}\sigma_w^2\mathbb{E}|\mu_{t-1-i}| = \beta^k\sigma_w^2\sum_{i=0}^{\infty}\beta^{2i}\left\{\mathbb{E}(\mu_{t-1-i}) + \mathbb{E}\left(|\mu_{t-1-i}| - \ \mu_{t-1-i}\right)\right\}\\ &= \beta^k\left(\frac{\sigma_w^2 E_r}{1-\beta^2}\right) + \beta^k\epsilon, \end{aligned} \tag{1.49}$$

where $\epsilon = \sigma_w^2\sum_{i=0}^{\infty}\beta^{2i}\mathbb{E}\left(|\mu_{t-1-i}| - \ \mu_{t-1-i}\right)$ represents the negligeable imperfection error (cf. Proposition 1 below). This leads to:

$$\mathrm{Cov}(r_{t+k},r_t) = \beta^{k-1}\left\{\beta\left(\frac{\sigma_w^2 E_r}{1-\beta^2}\right) + \sigma_{uw}\mathbb{E}(\sqrt{|\mu_{t-1}|})\right\} + \beta^k\epsilon. \tag{1.50}$$

Proposition 1 *Let the expected return process follow the equation,*

$$\mu_{t+1} = (1-\beta)E_r + \beta\mu_t + \sqrt{|\mu_t|}w_{t+1},$$

where $w_t \sim \mathcal{N}(0,\sigma_w)$, $E_r > 0$ and $\beta \in (0,1)$. The following result holds

$$\epsilon = \sigma_w^2\sum_{i=0}^{\infty}\beta^{2i}\mathbb{E}\left(|\mu_{t-1-i}| - \ \mu_{t-1-i}\right) \to 0$$

as $\Delta t \to 0$, with $n \to \infty$ and $n\Delta t \to \infty$, where the time index $t+k$ (in the time series notation) corresponds to the time $(n+k)\Delta t$, for $k \in \{0,1,2...\}$. Moreover, the term β^k factor of ϵ in (1.50) *further increases the convergence to zero.*

Proof:

$$\begin{aligned}\epsilon &= \sigma_w^2\sum_{i=0}^{\infty}\beta^{2i}\mathbb{E}\left(|\mu_{(n-1-i)\Delta t}| - \ \mu_{(n-1-i)\Delta t}\right) = -2\sigma_w^2\sum_{i=0}^{\infty}\beta^{2i}\mathbb{E}\left(\mu_{(n-1-i)\Delta t}\right)^-\\ &= -2\beta^{2n}\sigma^2\sum_{i=-\infty}^{n-1}\Delta t\beta^{-2(i+1)}\mathbb{E}\left(\mu_{i\Delta t}\right)^-,\end{aligned}$$

where we use the fact that $\sigma_w^2 = \sigma^2\Delta t$. This is of the form:

$$\kappa\mathbb{E}[\langle\nu^{n-1}, f\rangle],$$

where $f(x) = (x)^-$ and

$$\nu^m(\omega,dx) = \frac{1}{H^m}\sum_{i=-\infty}^{m}\beta^{-2(i+1)}\mathbb{I}_{\{\mu_{i\Delta t}(\omega)\in dx\}}$$

with

$$H^m = \sum_{i=-\infty}^{m}\beta^{-2(i+1)} = \beta^{-2(m+1)}\sum_{i=0}^{\infty}\beta^{2i} = \frac{\beta^{-2(m+1)}}{1-\beta^2} = \frac{\beta^{-2(m+1)}}{2\kappa\Delta t - \kappa^2\Delta t^2} \le \frac{\beta^{-2(m+1)}}{\kappa\Delta t},$$

where we have used the equality $\beta = (1 - \kappa\Delta t)$. Assume $\Delta t = \alpha(n) \to 0$, with $n \to +\infty$ and $n\Delta t \to +\infty$. Combining the weak convergence Theorem 2 of Pages, Panloup, et al. (2009) with the estimation of strong convergence of the symetrized Euler scheme using μ for X in Berkaoui et al. (2008), we get that for any continuous bounded function f

$$\frac{1}{H^n} \sum_{i=-\infty}^{n} \beta^{-2(i+1)} f(\mu_{i\Delta t}) \xrightarrow{n\to+\infty} \int_{\mathbb{R}+} f(x)\nu_0(dx) = 0, \ a.s.,$$

where ν_0 is the unique invariant measure of the CIR process (Gamma-type law). Due to the fact that the support of the function $f(x) = (x)^-$ is $\mathbb{R}^-$, and the support of a Gamma law is $(0, \infty)$, the integral and thus the limit of ϵ is equal to zero. □

On the other hand, assuming the positivity of the previous value of μ and removing the absolute value within the square root, the derivation of the autocovariance of returns is simpler and yields to:

$$\text{Cov}(r_{t+k}, r_t) = \beta^{k-1} \left\{ \beta \left(\frac{\sigma_w^2 E_r}{1 - \beta^2} \right) + \sigma_{uw} \mathbb{E}(\sqrt{\mu_{t-1}}) \right\}. \tag{1.51}$$

According to Proposition 1, the autocovariance of the process with $g(.) = \sqrt{|.|}$ boils down to equation (1.51) for $\mu_{t-1} \geq 0$, which is equation (1.11) in the text.

1.C The Kalman filter

As mentioned in the text, in this paper our analysis is performed without predictors, given that we focus on the implications arising from a modified interaction between past returns and expected returns. In this section we describe the (more general) procedure to estimate the unobservable process μ based on observations of both r and x, respectively the realized returns and predictor(s). The configuration investigated in the text corresponds simply to not consider the terms related to x in this section. In the next paragraph we present the algorithm for a state process $\{\mu_t\}$ with general dynamics described by equation (1.45) above. To do this, we use an extended version of the Kalman filter (see Anderson and Moore, 2012, Chap. 8), to which we add the predictors $\{x_t\}$ in order to present the estimation procedure of the full system including eventual predictors. The Kalman filter theory relies on the assumption that, conditional to the information available at time $t - 1$, denoted here D_{t-1}, i.e., $D_{t-1} = (r_1, x_1, r_2, x_2, ..., r_{t-1}, x_{t-1})$, the state variable μ_t has a Gaussian distribution. This assumption must also hold conditioned on D_t. In the configuration described above, using the extended Kalman filter consists in linearizing the function g around our last estimation of μ, i.e., replacing the term $g(\mu_t)$ by $g\left(\mathbb{E}(\mu_t|D_t)\right)$ in our procedure.

1.C.1 The algorithm

Following Pastor and Stambaugh (2009), we use the following notations:

$$z_t = \begin{bmatrix} r_t \\ x_t \end{bmatrix}, \quad a_t = \mathbb{E}(\mu_t|D_{t-1}), \quad b_t = \mathbb{E}(\mu_t|D_t), \quad f_t = \mathbb{E}(z_t|D_{t-1}), \quad P_t = \text{Var}(\mu_t|D_{t-1}),$$
$$Q_t = \text{Var}(\mu_t|D_t), \quad R_t = \text{Var}(z_t|\mu_t, D_{t-1}), \quad S_t = \text{Var}(z_t|D_{t-1}), \quad G_t = \text{Cov}(z_t, \mu_t|D_{t-1}).$$

Initialization We assume conditioning on the (unknown) parameters even if not explicitly specified and that D_0 denotes the null information.
Assuming that $\mu_1 \sim \mathcal{N}(E_r, V_\mu)$ and $r_1 \sim \mathcal{N}(E_r, V_r)$, given $V_x, V_{rx}, V_{r\mu}, V_{x\mu}$, we have first

$$a_1 = E_r, \quad P_1 = V_\mu, \quad f_1 = [E_r \ E_x]', \quad S_1 = \begin{bmatrix} V_r & V_{rx} \\ V_{rx} & V_x \end{bmatrix}, \quad G_1 = [V_{r\mu} \ V_{x\mu}]',$$
$$R_1 = S_1 - G_1 P_1^{-1} G_1',$$
$$Q_1 = P_1(P_1 + G_1' R_1^{-1} G_1)^{-1} P_1,$$
$$b_1 = a_1 + P_1(P_1 + G_1' R_1^{-1} G_1)^{-1} G_1' R_1^{-1}(z_1 - f_1).$$

Iteration We use the extended Kalman filter algorithm to derive, for $t = 2, ..., T$,

$$a_t = (1-\beta)E_r + \beta\mathbb{E}(\mu_{t-1}|D_{t-1}) + \mathbb{E}(g(b_{t-1})w_t|D_{t-1}) = (1-\beta)E_r + \beta b_{t-1}. \tag{1.52}$$
$$\begin{aligned} P_t &= \text{Var}((1-\beta)E_r + \beta\mu_{t-1} + g(b_{t-1})w_t|D_{t-1}) \\ &= \beta^2\text{Var}(\mu_{t-1}|D_{t-1}) + \text{Var}(g(b_{t-1})w_t|D_{t-1}) + 2\beta\text{Cov}(\mu_{t-1}, g(b_{t-1})w_t|D_{t-1}) \\ &= \beta^2 Q_{t-1} + g(b_{t-1})^2\sigma_w^2. \end{aligned} \tag{1.53}$$

We have:

$$S_t = \begin{bmatrix} \text{Var}(r_t|D_{t-1}) & \text{Cov}(x_t, r_t|D_{t-1}) \\ \text{Cov}(r_t, x_t|D_{t-1}) & \text{Var}(x_t|D_{t-1}) \end{bmatrix} = \begin{bmatrix} Q_{t-1} + \sigma_u^2 & \sigma_{uv} \\ \sigma_{vu} & \Sigma_{vv} \end{bmatrix} \tag{1.54}$$

$$G_t = \begin{bmatrix} G_t^1 \\ G_t^2 \end{bmatrix},$$

with

$$\begin{aligned} G_t^1 &= \text{Cov}(\mu_{t-1} + u_t, (1-\beta)E_r + \beta\mu_{t-1} + g(b_{t-1})w_t|D_{t-1}) \\ &= \beta Q_{t-1} + \text{Cov}(\mu_{t-1}, g(b_{t-1})w_t|D_{t-1}) \\ &+ \beta\text{Cov}(u_t, \mu_{t-1}|D_{t-1}) + \text{Cov}(u_t, g(b_{t-1})w_t|D_{t-1}) \\ &= \beta Q_{t-1} + g(b_{t-1})\sigma_{uw}, \end{aligned}$$

and

$$\begin{aligned} G_t^2 &= \text{Cov}((I-A)E_x + Ax_{t-1} + v_t, (1-\beta)E_r + \beta\mu_{t-1} + g(b_{t-1})w_t | D_{t-1}) \\ &= \text{Cov}(v_t, g(b_{t-1})w_t | D_{t-1}) \\ &= g(b_{t-1})\sigma_{vw}. \end{aligned}$$

Finally,

$$G_t = \begin{bmatrix} \beta Q_{t-1} + g(b_{t-1})\sigma_{uw} \\ g(b_{t-1})\sigma_{vw} \end{bmatrix}. \tag{1.55}$$

The last terms are functions of those previously computed:

$$R_t = S_t - G_t P_t^{-1} G_t', \tag{1.56}$$

$$Q_t = P_t(P_t + G_t' R_t^{-1} G_t)^{-1} P_t, \tag{1.57}$$

$$f_t = \begin{bmatrix} \mathbb{E}(\mu_{t-1}|D_{t-1}) \\ (I-A)E_x + Ax_{t-1} \end{bmatrix} = \begin{bmatrix} b_{t-1} \\ (I-A)E_x + Ax_{t-1} \end{bmatrix}. \tag{1.58}$$

The filtering term b_t is given by

$$b_t = a_t + P_t(P_t + G_t' R_t^{-1} G_t)^{-1} G_t' R_t^{-1}(z_t - f_t) = a_t + G_t' S_t^{-1}(z_t - f_t). \tag{1.59}$$

Following the notation in Pástor and Stambaugh (2009), denote

$$\begin{aligned} [m_t \; n_t'] &= P_t(P_t + G_t' R_t^{-1} G_t)^{-1} G_t' R_t^{-1} = G_t' S_t^{-1} \\ &= \text{Cov}(z_t', \mu_t | D_{t-1})[\text{Var}(z_t | D_{t-1})]^{-1} \\ &= [\beta Q_{t-1} + g(b_{t-1})\sigma_{uw} \quad g(b_{t-1})\sigma_{vw}] \begin{bmatrix} Q_{t-1} + \sigma_u^2 & \sigma_{uv} \\ \sigma_{vu} & \Sigma_{vv} \end{bmatrix}^{-1}. \end{aligned} \tag{1.60}$$

Notice that the terms defining m_t (and n_t) depend on $g(b_{t-1})$, which suggests a higher time dependence with respect to the terms in Pástor and Stambaugh (2009) setting. This implies that the level and change effect might have a more variable relative importance over time in our setting relative to the AR system.

From equation (1.59), we derive

$$b_t = a_t + [m_t \; n_t'](z_t - f_t) \tag{1.61}$$

$$= (1-\beta)E_r + \beta b_{t-1} + [m_t \; n_t'] \begin{bmatrix} r_t - b_{t-1} \\ x_t - (I-A)E_x - Ax_{t-1} \end{bmatrix} \tag{1.62}$$

$$= (1-\beta)E_r + (\beta - m_t)b_{t-1} + m_t r_t + n_t' v_t. \tag{1.63}$$

By repeated substitutions of the lagged values of $(b_t - E_r)$ in equation (1.63) we obtain:

$$b_t = E_r + \sum_{s=1}^{t} \lambda_s(r_s - b_{s-1}) + \sum_{s=1}^{t} \phi_s' v_s, \tag{1.64}$$

where $\lambda_s = m_s\beta^{t-s}$ and $\phi_s = n_s\beta^{t-s}$ and $(r_s - b_{s-1}) = r_t - \mathbb{E}(r_t|D_{t-1})$ is the forecast error. Equation (1.64) has the same structure than the equivalent expression in the AR predictive system of Pástor and Stambaugh (2009), but the coefficients m_s and n_s are modified, leading to different predictions. Equation (1.64) can be rewritten as a function of past returns instead of past forecast errors as follows

$$b_t = E_r + \sum_{s=1}^{t} \omega_s(r_s - E_r) + \sum_{s=1}^{t} \delta_s' v_s, \tag{1.65}$$

where,

$$\omega_s = \begin{cases} (\beta - m_t)(\beta - m_{t-1}) \dots (\beta - m_{s+1}) m_s & , \text{ for } s < t \\ m_s & , \text{ for } s = t. \end{cases}$$

and

$$\delta_s = \begin{cases} (\beta - m_t)(\beta - m_{t-1}) \dots (\beta - m_{s+1}) n_s & , \text{ for } s < t \\ n_s & , \text{ for } s = t. \end{cases}$$

If E_r is replaced by the sample mean in equation (1.65), it can be shown that the estimate of b_t is

$$\hat{b}_t = \sum_{s=1}^{t} \kappa_s r_s + \sum_{s=1}^{t} \delta_s' v_s, \tag{1.66}$$

where

$$\kappa_s = \frac{1}{t}\left(1 - \sum_{l=1}^{t} \omega_l\right) + \omega_s, \tag{1.67}$$

and $\sum_{s=1}^{t} \kappa_s = 1$. This expression has the same form than in the AR predictive system of Pástor and Stambaugh (2009), but the ω_s are functions of m_s, which has a different expression in our setting that depends on the level of μ_s. To see this, develop equation (1.63), add and subtract $m_t E_r$, rearrange terms and do backward substitution of $(b_t - E_r)$.

1.C.2 Steady state

Important results can be obtained assuming the system reach a steady state on the long run. Note that the results of Pages et al. (2009) ensure the existence of a steady state in the CIR discretization case, i.e., $g(.) = \sqrt{|.|}$. The expressions of the different elements defined at the beginning of section 1.C can be derived, at the equilibrium, removing the subscripts t and $t-1$ of equations (1.53) to (1.57). We obtain the following system to

solve for Q, the steady-state value of Q_t:

$$P = \beta^2 Q + g(b)^2\sigma_w^2, \quad S = \begin{bmatrix} Q + \sigma_u^2 & \sigma_{uv} \\ \sigma_{vu} & \Sigma_{vv} \end{bmatrix},$$
$$G = \begin{bmatrix} \beta Q + g(b)\sigma_{uw} \\ g(b)\sigma_{vw} \end{bmatrix}, \quad R = S - GP^{-1}G',$$
$$Q = P(P + G'R^{-1}G)^{-1}P.$$

After rearranging terms, this gives us the following quadratic equation for Q:

$$Q^2 + \xi_1 Q + \xi_2 = 0, \tag{1.68}$$

where

$$\begin{aligned}\xi_1 &= (1-\beta^2)(\sigma_u^2 - \sigma_{uv}\Sigma_{vv}^{-1}\sigma_{vu}) + 2g(b)\beta(\sigma_{uw} - \sigma_{wv}\Sigma_{vv}^{-1}\sigma_{vu}) - g(b)^2(\sigma_w^2 - \sigma_{wv}\Sigma_{vv}^{-1}\sigma_{vw}) \\ &= (1-\beta^2)\mathrm{Var}(u \mid v) + 2g(b)\beta\mathrm{Cov}(u, w \mid v) - g(b)^2\mathrm{Var}(w \mid v),\end{aligned} \tag{1.69}$$

and

$$\begin{aligned}\xi_2 &= g(b)^2\left((\sigma_{uw} - \sigma_{wv}\Sigma_{vv}^{-1}\sigma_{vu})^2 - (\sigma_u^2 - \sigma_{uv}\Sigma_{vv}^{-1}\sigma_{vu})(\sigma_w^2 - \sigma_{wv}\Sigma_{vv}^{-1}\sigma_{vw})\right) \\ &= g(b)^2\left(\mathrm{Cov}(u, w \mid v)^2 - \mathrm{Var}(u \mid v)\mathrm{Var}(w \mid v)\right).\end{aligned} \tag{1.70}$$

The solution is thus the positive root of (1.68):

$$Q = \frac{\sqrt{\xi_1^2 - 4\xi_2} - \xi_1}{2}. \tag{1.71}$$

Moreover, using the value of Q given by (1.71) and equation (1.60), we obtain the steady-state expressions of m and n:

$$m = (\beta Q + g(b)\mathrm{Cov}(u, w \mid v))(Q + \mathrm{Var}(u \mid v))^{-1}, \tag{1.72}$$
$$n' = (g(b)\sigma_{wv} - m\sigma_{uv})\Sigma_{vv}^{-1}. \tag{1.73}$$

1.D Bayesian procedure

This section describes the Bayesian analysis of the CIR predictive system. As in Appendix 1.C, we provide a description of the procedure for the full system, i.e., with eventual predictors. As Pástor and Stambaugh (2009), we use an MCMC procedure to obtain the posterior distribution of μ and θ the set of parameters, based on D the data available to the investor. We alternate between drawing μ from the posterior distribution $p(\mu|\theta, D)$ and drawing the parameters θ from the posterior $p(\theta|\mu, D)$.

1.D.1 Drawing μ_t

Given a set of parameters, we draw the time series of $\{\mu_t\}$ using the forward filtering, backward sampling approach of Carter and Kohn (1994) and Frühwirth-Schnatter (1994). The first stage consists in applying the Kalman filter procedure described above in section 1.C, with the current set of parameters. The sampling stage is the same as described in section B3.2. of the internet appendix of Pástor and Stambaugh (2009). However, given that one of the motivations to use a modified version of the AR predictive system is the belief that $\mu_t > 0$, we choose to use a rejection sampling methodology, i.e., we impose to each draw of μ_t to be positive at all times. Our procedure is the following, at each time step μ_t is drawn using the distribution $\mu_t|\mu_{t+1}, D_t$ which is Gaussian (due to the use of the extended Kalman filter) and thus can lead to eventual negative values (though rare, we have to consider this eventuality due to the discretization imperfection). In the case where a negative value is drawn at a specific time step t^* for μ_{t^*}, we redraw μ_{t^*} using the same distribution until a positive values is obtained, i.e., we draw μ_t using an acceptance-rejection method[23].

1.D.2 Prior distributions

As mentioned in the text, the priors used in the case of the CIR system are very similar to those in the AR system used by Pástor and Stambaugh (2009). Thus, we refer to its internet appendix for a detailed description of the prior distributions. A brief description, and a discussion of the slight modifications of E_r prior due to the specification of the CIR system, are done in sections 1.3.1 and 1.3.3 of the text.

1.D.3 Posterior distributions

Conditional on the current draw of $\{\mu_t\}$, the posterior distributions of the parameters are the same as describe in section B5.2. of the internet appendix of Pástor and Stambaugh (2009), except the terms affected by the new dynamics of μ described below, the others remain unchanged. In this section we denote as K the number of predictors (notice that in the results presented in the text $K = 0$ as no predictors are used) and T the last period at which returns are available for the estimation period of concern. We use $E_{x\mu_0}$ and $V_{x\mu_0}$ as notations for the prior mean and variance of the vector $[E_x \; E_r]'$. Let $\Sigma_{(vw)} = \begin{bmatrix} \Sigma_{vv} & \sigma_{vw} \\ \sigma_{wv} & \sigma_w^2 \end{bmatrix}$ and $q_t = \begin{bmatrix} x_t \\ \mu_t \end{bmatrix}$ for $t \in \{1, 2, ..., T\}$.

[23] In the case where after a maximum number of 500 trials μ_{t^*} is still negative we reject the set of parameters and the current draw of μ. We draw a new set of parameters to sample a whole new time series of μ_t. The percentage of parameters rejection is relatively small in all configurations we tested.

Posterior of E_x and E_r The posterior for $E_{x\mu} = [E_x \; E_r]'$ is still normal $E_{x\mu} \sim \mathcal{N}(\tilde{E}_{x\mu}, \tilde{V}_{x\mu})$ but with[24]

$$\tilde{V}_{x\mu} = \left(V_{x\mu_0}^{-1} + \sum_{t=1}^{T-1} L_2' \Sigma_{(vw)}^{g(\mu_t)} L_2 \right)^{-1} \quad \text{and} \quad \tilde{E}_{x\mu} = \tilde{V}_{x\mu} \left(V_{x\mu_0}^{-1} E_{x\mu_0} + L_2' \sum_{t=1}^{T-1} \Sigma_{(vw)}^{g(\mu_t)} \left(q_{t+1} - L_1 q_t \right) \right),$$

where $L_1 = \begin{bmatrix} A & 0 \\ 0 & \beta \end{bmatrix}$, $L_2 = \begin{bmatrix} I_K - A & 0 \\ 0 & 1 - \beta \end{bmatrix}$ and $\Sigma_{(vw)}^{g(\mu_t)} = \Sigma_{(vw)} \circ \begin{bmatrix} \mathbf{1}_{K\times K} & g(\mu_t)\mathbf{1}_{K\times 1} \\ g(\mu_t)\mathbf{1}_{1\times K} & g(\mu_t)^2 \end{bmatrix}$.

Posterior of A and β The posterior of $\mathbf{b} = [\text{vec}(A') \; \beta]'$ is still the Gaussian distribution described in section B.5.2.1. of internet appendix of Pástor and Stambaugh (2009). Using the same notation we have:

$$z = Z\mathbf{b} + \text{errors},$$

but the covariance matrix of the error terms is instead

$$\left(\Sigma_{(vw)} \otimes I_{T-1} \right) \circ \Lambda_{1,T-1},$$

where $\Lambda_{1,T-1}$ is the following $[(T-1)(K+1)] \times [(T-1)(K+1)]$ matrix:

$$\begin{bmatrix} & & & \begin{bmatrix} g(\mu_1) & & \mathbf{1} \\ & \ddots & \\ \mathbf{1} & & g(\mu_{T-1}) \end{bmatrix} \\ & \mathbf{1}_{K(T-1)\times K(T-1)} & & \vdots \\ & & & \begin{bmatrix} g(\mu_1) & & \mathbf{1} \\ & \ddots & \\ \mathbf{1} & & g(\mu_{T-1}) \end{bmatrix} \\ \begin{bmatrix} g(\mu_1) & & \mathbf{1} \\ & \ddots & \\ \mathbf{1} & & g(\mu_{T-1}) \end{bmatrix} & \cdots & \begin{bmatrix} g(\mu_1) & & \mathbf{1} \\ & \ddots & \\ \mathbf{1} & & g(\mu_{T-1}) \end{bmatrix} & \begin{bmatrix} g(\mu_1)^2 & & \mathbf{1} \\ & \ddots & \\ \mathbf{1} & & g(\mu_{T-1})^2 \end{bmatrix} \end{bmatrix},$$

which is equal to the matrix of the AR system except for the terms involving μ (i.e., bottom and right block of the matrix above).

Posterior of Σ We apply here the decomposition of Pástor and Stambaugh (2009) by changing variables from $\Sigma = \begin{bmatrix} \sigma_u^2 & \sigma_{uv} & \sigma_{uw} \\ \sigma_{vu} & \Sigma_{vv} & \sigma_{vw} \\ \sigma_{wu} & \sigma_{wv} & \sigma_w^2 \end{bmatrix}$ to the set of (Σ_{11}, C, Ω), where $\Sigma_{11} = \begin{bmatrix} \sigma_u^2 & \sigma_{uw} \\ \sigma_{wu} & \sigma_w^2 \end{bmatrix}$ and C and Ω are the slope and the residual covariance matrix of the regression of v on (u, w). The procedure to draw (Σ_{11}, C, Ω) remains the same, except that given a draw of the time series of $\{\mu_t\}$, and conditional on (E_x, A, E_r, β), the sample of

[24] The operator $\circ$ denotes the Hadamard product, also known as the element-wise product. I_K is the identity matrix with dimension $K \times K$. For integers m and n, $\mathbf{1}_{m\times n}$ denotes a matrix with m rows and n columns whose all entries are 1.

residuals of $\{w_t, t = 2, ..., T\}$ is no longer the time series $\{\mu_t - (1-\beta)E_r - \beta\mu_{t-1}, t = 2, ..., T\}$ but $\left\{\frac{\mu_t - (1-\beta)E_r - \beta\mu_{t-1}}{g(\mu_{t-1})}, t = 2, ..., T\right\}$. Thus using $g(.) = \sqrt{|.|}$ can lead to very large observations of w in magnitude when the current value of μ_{t-1} is close to zero. Thereby to avoid using biased estimators of variances and covariances involving w, we use a robust estimator to measure these terms. Specifically, let X denote the $(T-1) \times 2$ matrix of $[u_t, w_t]$ for $t = 2, ..., T$. Instead of using the classical estimator of variance for Σ_{11}: $\hat{\Sigma}_{11} = \frac{1}{T-1}(X'X)$, we use the robust and widely used *Minimum Covariance Determinant Estimator* (Fast MCD) of Rousseeuw and Driessen (1999) to compute $\hat{\Sigma}_{11}$. This method gives a weight vector ω of size $(T-1) \times 1$ with entries 0 or 1 for each observation based on a Mahalanobis distance criterium. Hence, in order to compute the parameters of the regression of v on (u, w) and derive the posterior distributions parameters of (Σ_{11}, C, Ω) described in section B5.2.2. of the internet appendix of Pástor and Stambaugh (2009), we use the corresponding reweighted vector $[\omega\ \omega] \circ X$ instead of X.

Chapter 2

Stock Market Volatility Dynamics: A Volume Filtered-GARCH Model

We present a two-factor volatility model to study the impact of news arrival and trading volume on stock returns variance. The model can explicitly account for the association between volatility and volume, as well as the persistence in equity variance. Unlike the standard "Mixture of Distribution Hypothesis", the conditional variance is governed by the stochastic information arrival and adds a persistent GARCH component, in order to disentangle transient from persistent volatility variations. The common observation that large volumes are associated with high volatility is explained by the fact that unexpected shocks in volume increase volatility, which is not the case for expected volumes. Furthermore, the persistence of volatility is essentially unrelated to volume implying that the latter does not explain ARCH effect. Finally, we find that unexpected shocks in trading volume and the persistent GARCH component are both main drivers of volatility dynamics.

Keywords: Mixture of Distribution, stochastic volatility, GARCH, trading volume.

2.1 Introduction

The financial literature has documented a number of theoretical arguments to explain the empirical relationship between stock price movements and trading volumes (see Karpoff, 1987, for a survey). In particular, mixture of distribution models predict that both trading activity and equity volatility are closely linked to the latent information arrival rate, implying a joint dynamics for these variables (see Clark, 1973; Epps and Epps, 1976; Tauchen and Pitts, 1983; Andersen, 1996). These specifications provide a theoretical framework justifying the empirical evidence that large price changes are accompanied by high volume. Hence, volume is likely to be related to short-term fluctuations of volatility. However, on the other hand, these theoretical specifications cannot jointly capture major stylized features of volatility and stock returns (see Gallant, Rossi, and Tauchen, 1992). In particular, several studies highlight that mixture models cannot account for the pervasive persistence in stock price changes variance (see Lamoureux and Lastrapes, 1994; Liesenfeld, 1998; Watanabe, 2000). In the same vein, Andersen, Bollerslev, and Diebold (2007) suggest that separating the jump moves of volatility associated with specific news announcements from the smooth continuous moves results in significant volatility forecast improvements. In addition, recent studies shedding light on the link between stock variance and economic activity also suggest that stock market volatility can be governed by two components, one that is slowly varying (see for instance Engle and Rangel, 2008; Engle, Ghysels, and Sohn, 2013). From a perspective focusing on news arrival and trading volume, this paper proposes a model to disentangle the sources of changing variance dynamics, in order to reconcile the empirical properties of volatility with theoretical models incorporating volume.

We introduce a parametric framework for the return generating process and we model endogenously its interactions with trading volume. The proposed specification implies a two-factor volatility model, governed by the stochastic information flow modeled by a non-linear function of volume, as well as an additional persistent mean-reverting GARCH component unrelated to trading activity. This study complements the existing literature on the relationship between price changes and volume, with an informative parametric model where some fraction of the persistence in volatility can be explicitly unrelated to the autocorrelation of information arrival. Therefore, the model captures both strong empirical evidences that volatility is persistent and is related to volume.

The strength of the relationship between stock price variations and trading volume is evidenced by several papers. In particular, Hiemstra and Jones (1994) document a non-linear causality between returns and volume. Campbell, Grossman, and Wang (1993) and Avramov, Chordia, and Goyal (2006) study return reversals caused by high volume[1], while Wang (1994), Cooper (1999), and Llorente, Michaely, Saar, and Wang (2002) examine price continuation following large volume, depending on information asymmetry. Our paper concentrates on the impact of volumes on stock return heteroskedasticity. Regarding this specific issue, an old and pervasive observation at the origin of several models is

[1]Especially for illiquid stocks, see for instance Pastor and Stambaugh (2003), Acharya and Pedersen (2005), and Da, Liu, and Schaumburg (2013) on return reversals and liquidity.

the positive association between stock variance and trading volume. In particular, this empirical regularity is highlighted by Gallant et al. (1992) using a seminonparametric approach for modeling price changes and volatility. It is further motivated by differences of opinions and expectations among traders by Harris and Raviv (1993) and Shalen (1993).

Our model is built on two branches of the literature: first, the source of the volatility-volume relationship, and second, recent advances regarding two-factor volatility models. The studies of Clark (1973); Epps and Epps (1976); Tauchen and Pitts (1983); Andersen (1996) theoretically justify the daily association between volatility and volume using a Mixture of Distribution Hypothesis[2] (MDH). The latter implies that the conditional variance as well as trading volume are functions of the latent rate of information that flows into the market. Moreover, empirical studies such as Lamoureux and Lastrapes (1990a) and Fleming, Kirby, and Ostdiek (2006), further suggest that volume captures a large non-persistent component of return volatility (an effect absent in ARCH models[3]). The resulting specifications of the Mixture of Distribution Hypothesis differ from the widespread ARCH-type models and imply rather a stochastic volatility representation driven by the information flow. Note that if the latter is positively autocorrelated across days, this specification should naturally generates autocorrelation in the magnitude of price changes, and capture volatility persistence as GARCH specifications[4]. However, many papers such as Lamoureux and Lastrapes (1994); Liesenfeld (1998); Watanabe (2000) find that the MDH framework fails to capture the persistence exhibited by price change variance, suggesting that the latter can be unrelated to the autocorrelation of information arrival[5]. Richardson and Smith (1994) also support alternative theories that do not rely solely on information arrival. In addition, standard MDH models cannot account for jumps in volatility. In particular, the flexible framework of our paper combines advantages of both MDH and GARCH models in order to disentangle the factors driving volatility variations. We propose a model where the return variance is governed by the arrival of new information, i.e., the stochastic mixing variable, as well as a smoother mean-reverting component unrelated to information flow[6], in order to capture volatility persistence. The existence of this second component is also motivated empirically by Andersen et al. (2007), as well as Fleming et al. (2006) suggesting a model allowing persistence in return volatility unrelated to persistence in trading volume. Furthermore, the results of Engle and Lee (1999); Engle and Rangel (2008); Colacito, Engle, and Ghysels (2011); Engle and Sokalska (2012); Engle et al. (2013), studying component-GARCH models, underline the significant role of a slow-moving component in volatility.

Our so-called Volume Filtered-GARCH (VF-GARCH) model decomposes stock return

[2]Foster and Viswanathan (1995) also develop a speculative trading model they apply with 30-minute volume and price data.

[3]See Engle (1982); Bollerslev (1986); Engle (1990); Nelson (1991); Glosten, Jagannathan, and Runkle (1993) on ARCH models and principal variations.

[4]Nelson (1990) shows that the discrete version of the continuous time ARCH model can be expressed as an MDH model.

[5]Darolles, Le Fol, and Mero (2015) also develop an extended version of the MDH that can explicitly account for liquidity frictions in volume.

[6]Former studies proposing two-factor volatility models include Ding and Granger (1996); Gallant, Hsu, and Tauchen (1999); Alizadeh, Brandt, and Diebold (2002); Chernov, Gallant, Ghysels, and Tauchen (2003); Maheu and McCurdy (2004); Andersen et al. (2007); Adrian and Rosenberg (2008).

variance as the product of two terms: 1) a stochastic mixing variable, referred as volume component, and 2) a predictable filtered component modeled as a GARCH-type model[7]. This specification implies: the empirically motivated association between volatility and volume, the persistent mean-reverting behavior for conditional variance, and an asymmetric and leptokurtic normal log-normal conditional distribution for returns. Specifically, we do not model the latent news process directly, but rather decompose the mixing variable as a combination of expected and unexpected change in volume, i.e., the predictable part and the innovation of the volume process. This splitting allows to account for the volume generating process to avoid simultaneity bias. Statistical tests reject simpler versions of the model. Fleming et al. (2006) also show that the unpredictable component of volatility is closely linked to the innovation in contemporaneous trading volume, i.e., a portion of the information flow is unanticipated news. In this way, our approach allows an informative interpretation of the role of both expected and unexpected news in the conditional variance.

Applying a two-step Maximum Likelihood method to estimate our specification, a main finding is that the two components of the mixing variable, that are expected and unexpected volume, have different impacts on return variance. More precisely, unlike Bessembinder and Seguin (1993), we find that expected volume of trading brings down volatility, while unexpected shocks in volume increases it. In other words, if a large change in information flow, i.e., trading activity, is anticipated the volatility is reduced. On the other hand, if a large unanticipated change occurs, the volatility is increased. Furthermore, an important feature of the model is that the correlation between contemporaneous innovations in volume and return is a key determinant of the distribution of stock returns, implying excess kurtosis and non-zero skewness, and characterizes the joint return/volume distribution. This parameter also controls for the risk-return tradeoff in equities. Indeed, if one assumes that shocks in volume are a good proxy for news shocks then this correlation is an informative indication on how stock price changes are affected by new information.

Through the GARCH component, our parametric approach also captures the persistence and leverage effect in volatility, and allows to investigate the impact of volume on these usual stylized features of stock variance. After estimating the model, we do not find evidence that the volume component annihilates ARCH effect: the estimated parameters of persistence in our model remain consistent with standard GARCH estimates. Thus, as suggested by Fleming et al. (2006), we find that the persistence of variance is mainly unrelated to trading volume. Moreover, the well documented leverage effect also remains significant, meaning that return innovations have an asymmetric impact on volatility independently of trading volume.

Furthermore, for most stocks, the model reveals that the volume component contributes to a large part of variations in stock variance. More specifically, the component affected by contemporaneous unexpected shocks in volume appears to clearly drive volatility. This finding also implies that a large portion of daily variance is a priori unpredictable

[7]We use the asymmetric GARCH model of Glosten et al. (1993).

as it is due to the innovation in volume, independent of previous information available. On the other hand, most of the (substantial) remaining fraction of variations in stock variance is associated with the persistent GARCH component and is therefore unrelated to trading volume, highlighting the limits of standard MDH models to capture the full dynamics of equity volatility.

Finally, our specification also provides an informative interpretation of volume and volatility variations during important news events such as earning surprises. Using the example of IBM, we find that large price changes occurring during these special announcements are associated with large unexpected volume. This implies large but short-lived increases in volatility estimates of our model, attributed to the volume component. Our results corroborate the findings of Bajgrowicz, Scaillet, and Treccani (2015) that the arrival of news induces bursts of volatilities.

The rest of the paper is organized as follows. Section 2.2 introduces the approach to model stock returns and interactions with trading volume. Section 2.3 examines the properties of the distribution of returns implied by the model. Section 3.7 presents our main empirical analysis, while section 2.5 focuses on empirical regularities over sub-samples. Finally section 2.6 discusses the impact on volatility of trading volumes occurring during important news events.

2.2 A two-component volatility model

This section describes the proposed framework to explore the specific issue of the relationship between stock price changes and trading volume of individual stock returns. The model is parametric and is based on a variant of the Mixture of Distribution Hypothesis (MDH) implying two components for the stock variance dynamics. The equation of the daily return r_t is given by:

$$r_t = \mu + \epsilon_t, \tag{2.1}$$

where μ is a constant. To motivate the specification of this paper, let δ_t^i denotes the i^{th} intraday price increment on day t, and g_t the stochastic rate of information arrival. Following previous studies using MDH models[8], the return innovation is decomposed as follows:

$$\epsilon_t = \sum_{i=1}^{g_t} \delta_t^i. \tag{2.2}$$

In a liquid market, the quantity g_t is therefore large and time-varying. Furthermore, for a given day t the intraday price increments δ_t^i are assumed to be i.i.d. across i:

$$\delta_t^i \overset{i.i.d}{\sim} (0, \sigma_t^2). \tag{2.3}$$

Unlike typical MDH models, σ_t the variance of these small price increments is not assumed constant over time. Therefore the volatility of the intraday price increments on day t is allowed to differ from one day to another, adding potentially a source of persistence in

[8]See Clark (1973), Tauchen and Pitts (1983), Lamoureux and Lastrapes (1990a), Andersen (1996).

the conditional variance unrelated to g_t.

Let $\mathcal{F}_{t-1}$ denote the information set available at time $t-1$ and assume σ_t is $\mathcal{F}_{t-1}$-measurable. The original Central Limit Theorem does not directly apply in (2.2) because g_t is stochastic. However if g_t is sufficiently large, using Clark (1973)'s Theorem 3, under weak regularity conditions we have[9]:

$$\epsilon_t \mid g_t, \mathcal{F}_{t-1} \sim \mathcal{N}(0, \sigma_t^2 g_t). \tag{2.4}$$

It follows directly that $\mathrm{Var}(\epsilon_t \mid g_t, \mathcal{F}_{t-1}) = \mathbb{E}[\epsilon_t^2 \mid g_t, \mathcal{F}_{t-1}] = \sigma_t^2 g_t$. Therefore, this observation suggests a parametric representation of the following form for ϵ_t:

$$\epsilon_t = \sqrt{\sigma_t^2 g_t} z_t, \tag{2.5}$$

where z_t is a Gaussian random variable, with zero mean and unit variance. z_t is hereafter assumed to be standard Gaussian white noise $z_t \sim \mathcal{N}(0,1)$, i.i.d across t.

We can make an analogy between this specification and the models studied in the component-GARCH literature: see Engle and Lee (1999); Engle and Rangel (2008); Engle et al. (2013). These papers are focused on the decomposition of stock variance between a short-term and a long-term component, and they also implies a representation in the form of a product for the conditional volatility.

Introducing σ, the benchmark unconditional volatility of intraday price variations, we redefine σ_t as:

$$\sigma_t^2 \;=\; \sigma^2 \tau_t, \tag{2.6}$$

where τ_t denotes the intensity of variations relative to the benchmark value, and has unit mean. Intuitively, τ_t should be mean-reverting and persistent. Finally the equation for the daily return is the following:

$$r_t = \mu + \sigma \sqrt{\tau_t \times g_t} z_t. \tag{2.7}$$

A main consequence is that our model implies two time-varying components for the daily variance of return: 1) τ_t related to the range of variations of price increments, and 2) g_t related to the information arrival rate. The motivation being to investigate the different features of the conditional volatility of returns, the specification described in the next sections will allow persistence in both τ_t and g_t. Indeed, if τ_t and/or g_t is serially correlated, this naturally generates autocorrelation in return variance. The process τ_t will be specified in order to capture the persistent component of volatility that is unrelated to the information arrival rate, whereas g_t will account for the impact of news arrival via trading volume on return's variance.

The two-factor decomposition of equation (2.7) differs from the Volume Augmented-GARCH (VA-GARCH) of Lamoureux and Lastrapes (1990a), directly adding a volume term in the GARCH variance equation: $h_t = \omega + \alpha_1 (r_{t-1} - \mu)^2 + \beta_1 h_{t-1} + \gamma V_t$, where h_t is the conditional stock variance and the volume V_t is considered as weakly exogenous. As

[9] cf., Clark (1973) and Andersen (1996) for a discussion.

noted by Fleming et al. (2006) and Fleming, Kirby, and Ostdiek (2005), this specification constraints the volume effect to decay at the same rate as ARCH effect, i.e., the coefficients on V_{t-s} and r^2_{t-s-1} are both proportional to $(\beta_1)^s$ for $s > 0$. In particular, the specification described in the next sections will allow to overcome this issue.

2.2.1 The mixing variable as stochastic volume component

In equation (2.7), g_t represents the component of stock variance directly related to the information arrival rate. This component is usually called the mixing variable. Several research studies highlight the economic and empirical link between trading volume and volatility. The usual evidence is that high volume tends to increase volatility[10] as a result of information arrival. In the variance specification of equation (2.7), the day-to-day volume impact is integrated through the mixing variable g_t if a stock is heavily traded (in that case volume relates directly to the information arrival process). Therefore, the volume information is incorporated into a specific component of the model, and given this separation the dynamics of τ_t presented in section 2.2.2 is expected to be adjusted from volume effect and to only capture remaining volatility persistence unrelated to trading activity. For this reason, our model is referred as Volume Filtered-GARCH (VF-GARCH).

As suggested by Clark (1973), Tauchen and Pitts (1983) and Andersen (1996), the positivity assumption of g_t makes the Poisson and the log-normal distributions natural candidates for the marginal distribution of the mixing variable. This paper focuses on the log-normal distribution as this choice results in a mixture of distributions for the return, able to replicate stylized features of realized stock returns in a tractable way. Indeed, as described below the return r_t, conditional to the information set available at time $t-1$, $\mathcal{F}_{t-1}$, will have a normal log-normal probability distribution, i.e., potentially asymmetric and leptokurtic. This specific mixture has been used and is well documented in the literature focused on the price-volume variations[11]. In particular, Clark (1973) models price changes and finds that the likelihood of a normal log-normal distribution is significantly higher than the likelihood for any stable distributions. Furthermore Tauchen and Pitts (1983) report that the likelihood of their model, for price fluctuations and volumes, is much higher when they use a log-normal distribution for the mixing variable. Richardson and Smith (1994) find that this distribution appears to be consistent with Dow Jones firms data. The normal log-normal mixture is also used to model innovations in returns in stochastic volatility models, (see for instance Ghysels, Harvey, and Renault, 1996).

The modeling of the mixing process, unobservable in practice, is a key issue. Our paper does not propose a latent variable approach, but rather introduces a specification for g_t related to the impact of different types of news, approximated by expected and unexpected volume of trading. Certain papers (such as Clark, 1973; Lamoureux and Lastrapes, 1990a) propose to use trading volume as a direct proxy for the mixing variable,

[10] For instance, Harris and Raviv (1993) and Shalen (1993) find that large volume of trading tends to announce large subsequent price variations in magnitude, that is, high volatility.

[11] Hsieh (1989) also uses the normal log-normal mixture to estimate ARCH and GARCH models for foreign currencies and finds that it fits the data quite well.

i.e., they assume that variations in trading volume capture exactly the amount of daily information flowing into the market. In that case, the model can be directly estimated by maximum likelihood. However, if volume and returns are jointly determined, introducing volume as an exogenous variable induces a simultaneity bias. In addition, raw volume may not perfectly reflect the amount of information, and studies such as Bessembinder and Seguin (1993) suggest that using directly volume may lead to a lack of flexibility in the model. Our approach consists in using a flexible function of volume as a proxy for the mixing variable, so that the model can be estimated directly by maximum likelihood, and also accounts for the volume generating process to avoid simultaneity bias.

We propose the following log-normal specification, implying that the mixing variable is a non-linear function of trading volume:

$$\log g_t = \gamma_1 v_t^e + \gamma_2 v_t^u. \tag{2.8}$$

where γ_1, γ_2 are constant parameters. v_t^e and v_t^u denote respectively *expected and unexpected change in the volume of trading*, i.e., the predictable component of variations in volume and the innovation[12]. As explained below, the latter is assumed to be Gaussian leading to a log-normal conditional distribution for g_t.

The decomposition of expected and unexpected volume is first motivated by Bessembinder and Seguin (1993) finding respectively that the magnitudes of the impact of anticipated and unanticipated volume shocks are different and asymmetric depending on their signs. Furthermore, Fleming et al. (2006) shows that a large non-persistent component of return volatility is related to the non-persistent component of volume. Therefore, the distinction of anticipated and unanticipated news (volume innovation) appears essential for exploring the different aspects of the return-volume relationship. Equation (2.8) allows a distinct impact of expected and unexpected information arrival through volume, and also an asymmetric effect depending on the sign of the shocks (due to the log). Note that we do not impose any restriction on the parameters involved in equation (2.8). If volume is positively correlated across days (due to the persistent number of information arrivals), this will be captured by the autoregressive structure of v_t^e (described below). Then, if this persistent part of volume is directly related to the daily variance of return, the parameter γ_1 should be statistically significant. Finally, as suggested by Fleming et al. (2006), γ_2 related to the innovation in volume, should be significant. Note that we can recover the initial MDH specification using trading volume as a direct proxy for the mixing variable from the one proposed in equation (2.8).

Using daily trading volume of individual component stocks of the S&P 100 Index[13], we define expected and unexpected volume processes using autoregressive-moving average (ARMA) models. Note that we do not decompose volume time series between a common (market) and a specific component, as discussed in Lo and Wang (2000) and Białkowski, Darolles, and Le Fol (2008) for instance. Given that we consider large capitalization stocks,

[12] An alternative decomposition of volume can be found in Andersen (1996), who proposes to split trading volume between an informed and a noise component, following both a Poisson distribution.

[13] Data used in the empirical section are further described in section 3.7.

i.e., heavily traded stocks, we do not consider an information-based and a liquidity-based component in volume as in Darolles et al. (2015). Indeed, our approach does not focus on disentangling market and stock-specific information, or liquidity-based versus informed activity, but rather expected and unexpected news through shocks in volume. Therefore we assume implicitly that expected and unexpected volumes relate directly to news arrival.

As volume time series V_t exhibit a high positive skewness, we model the logarithm of volume[14] in order to adjust for this asymmetry. As suggested by KPSS statistical test and empirical serial correlation, we postulate that the difference in $\log V_t$ time series (adjusted with dummy variables in order to control for the day of the week) follows an ARMA(p, q) process[15]. Hence log-volumes are modeled as follows:

$$\{1 - \phi_1 L - ... - \phi_p L^p\} (1 - L) \log V_t = \theta_0 + \{1 + \theta_1 L + ... + \theta_q L^q\} u_t, \tag{2.9}$$

or more compactly:

$$\{1 - \phi(L)\} (1 - L) \log V_t = \theta_0 + \{1 + \theta(L)\} u_t,$$

where L denotes the lag operator, and u_t are i.i.d across t with $u_t \sim \mathcal{N}(0, \sigma_u^2)$. $\phi(L)$ denotes the autoregressive structure, while $\theta(L)$ is the moving-average part. For each volume time series, the best model is selected according to Akaike Information criterion[16]. Selected models are presented in the right part of Table 2.3 and are further discussed in section 3.7. Note that the ARMA models' R^2 are about 26%, highlighting the relative adequacy of the models.

For each asset series, the *unexpected* (v_t^u) and *expected* (v_t^e) volume are formally defined as:

$$\begin{aligned} v_t^u &= u_t, \\ v_t^e &= \log \frac{V_t}{V_{t-1}} - u_t, \end{aligned}$$

implying that $\log \frac{V_t}{V_{t-1}} = v_t^e + v_t^u$. Therefore we have $\mathbb{E}[v_t^u | \mathcal{F}_{t-1}] = 0$, and $\mathbb{E}[v_t^e | \mathcal{F}_{t-1}] = v_t^e$, where $\mathcal{F}_{t-1}$ denotes the information set available at time $t - 1$. In the spirit of Fleming et al. (2006), the component involving v_t^u in g_t will capture the impact of the non-persistent component of volume on volatility. Given the levels of R^2 mentioned above, a model using volume, without any distinction between expected and unexpected components, would be presumably over influenced by the latter given the larger variability of innovations and despite the non-negligible part of variance explained by the predictable part of ARMA models. This further motivates the flexibility of equation (2.8) with two distinct parameters γ_1 and γ_2.

[14] The use of log-volume is common in the financial literature, see for instance Gallant et al. (1992); Campbell et al. (1993); Hiemstra and Jones (1994).

[15] Similarly, Bessembinder and Seguin (1993) propose to use an AR(10) model to decompose trading volume on future contracts.

[16] Given that most autocorrelations beyond lag 5 are insignificant, the orders of ARMA models are tested with $p, q \leq 5$.

Finally, to account for the potential joint dependence of returns and volumes, the shock in volume is allowed to be correlated with the contemporaneous shock in return (the innovation z_t), i.e., $\text{Corr}(z_t, u_t) = \text{Corr}(z_t, v_t^u) = \rho$, for all t. Indeed, empirical evidence can be found that there is a clear association in the timing of outliers in returns and volume (see Ying, 1966; Schwert, 1989; Gallant et al., 1992). The key parameter ρ, referred to as *the correlation between unexpected return and unexpected volume*, has a crucial importance in the return distribution (especially for conditional moments), as well as in the joint distribution return/volume, as shown in the section 2.3. Note that this parameter can be seen as an approximation of the correlation between news shocks and return shocks. The random vector $[z_t \; v_t^u]'$ is assumed to be distributed, i.i.d. across t, as follows:

$$\begin{bmatrix} z_t \\ v_t^u \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho\sigma_u \\ \rho\sigma_u & \sigma_u^2 \end{bmatrix} \right). \tag{2.10}$$

Using equations (2.7) and (2.8), the daily return can be expressed as:

$$r_t = \mu + \sigma\sqrt{\tau_t} \exp\left\{ \frac{\gamma_1 v_t^e}{2} \right\} \times \exp\left\{ \frac{\gamma_2 v_t^u}{2} \right\} z_t. \tag{2.11}$$

The daily random shock is therefore normal log-normal. A consequence of this distribution is that the potential departure from normality is due to the time-varying arrival rate, i.e., the volume innovation v_t^u in g_t. On the other hand, if the latter is constant over time ($v_t^u = 0$), then the stock return is conditionally Gaussian. This feature of the model is in line with the studies of Clark (1973) and Ané and Geman (2000), linking deviations from normality to the existence of variations in volume.

As discussed in details in section 2.3, the expectation of the normal log-normal shock in equation (2.11) is non zero. The daily return can be rewritten as:

$$r_t = \mu_t + \sigma\sqrt{\tau_t} \exp\left\{ \frac{\gamma_1 v_t^e}{2} \right\} \times \epsilon_t^r, \tag{2.12}$$

where

$$\mu_t = \mu + \sigma\sqrt{\tau_t} \exp\left\{ \frac{\gamma_1 v_t^e}{2} \right\} \mathbb{E}\left[\exp\left\{ \frac{\gamma_2 v_t^u}{2} \right\} z_t \right], \tag{2.13}$$

$$\epsilon_t^r = \exp\left\{ \frac{\gamma_2 v_t^u}{2} \right\} z_t - \mathbb{E}\left[\exp\left\{ \frac{\gamma_2 v_t^u}{2} \right\} z_t \right], \tag{2.14}$$

so that the expectation of the shock ϵ_t^r is zero.

2.2.2 An additional persistent factor in volatility

The existence of a second factor driving volatility is motivated by the failing of MDH specifications to capture volatility persistence. The component τ_t is introduced in order to describe variations of price changes unrelated to the information arrival rate. Following conclusions of Fleming et al. (2006), the model allows for a persistent component of

volatility that is linearly unrelated to the persistent component of volume, i.e., τ_t captures return variance's persistence that is independent of the short-term trading activity. Given the desired persistent mean-reverting behavior of the variable τ_t and the predictability assumption (cf., equation 2.4), this component is modeled using an asymmetric GJR-GARCH(1,1) variance process[17] with unit expectation:

$$\tau_t = \left(1 - \alpha_1 - \beta_1 - \frac{\alpha_2}{2}\right) + (\alpha_1 + \alpha_2 \mathbb{1}_{r_{t-1}<\mu_{t-1}}) \frac{(r_{t-1} - \mu_{t-1})^2}{\sigma^2 g_{t-1}} + \beta_1 \tau_{t-1}, \tag{2.15}$$

where $\mathbb{1}_{r_{t-1}<\mu_{t-1}}$ is an indicator function equal to 1 if $(r_{t-1} - \mu_{t-1}) < 0$ and 0 otherwise. The conditions on the parameters $\alpha_1, \alpha_2, \beta_1 \geq 0$ and $\alpha_1 + \beta_1 + \frac{\alpha_2}{2} < 1$ ensure the strict positivity of the component τ_t.

This specification relaxes the assumption of constant variance for intraday price increments, and captures a number of stylized facts regarding the volatility of stock prices: persistence, mean-reversion, and asymmetric impact of past return shocks. We allow explicitly the existence of persistence in volatility unrelated to the autocorrelation in g_t. Indeed, given the scaling by $\sigma^2 g_{t-1}$ in the second term of the right hand side of equation (2.15), this specification has the advantage to only capture the persistence in τ_t. Hence, it does not interact directly with the other component of variance, i.e., g_t, and therefore does not constrain the coefficient on the lagged squared returns to decay at the same rate as the lagged mixing variable, i.e., g_{t-1} (unlike the model studied in Lamoureux and Lastrapes, 1990a). As suggested by Fleming et al. (2006) and Fleming et al. (2005), the specification with two components described in equations (2.7) and (2.15) allows volume effect (described in the next section) to decay at a different rate than the ARCH effect captured by τ_t. Hence if the variance of price increments is persistent independently of the information flow and trading activity, the coefficients in the dynamic of τ_t will be presumably significant, even if volume does not provide much information about stock variance.

2.3 Properties of the return distribution

In this section, we analyze the distribution of stock returns implied by the model previously introduced. In particular, we focus on the moments of the distribution and the role of the different parameters.

Given equation (2.11) and the fact that τ_t and v_t^e are determined given $\mathcal{F}_{t-1}$, the conditional returns follow a normal log-normal (NLN) distribution. Because of the potentially non-zero correlation between unexpected return z_t and unexpected volume v_t^u, i.e., ρ in equation (2.10), the properties of the NLN conditional distribution generate a variety of patterns for conditional moments of returns as well as moments of the joint (return, volume) distribution. The correlation ρ will be the key driver of many aspects of the distribution.

As noted in Yang (2008), the first four centered moments of the NLN product, i.e.,

[17] See Glosten et al. (1993).

$\exp\{\frac{\gamma_2 v_t^u}{2}\} z_t$ are known, implying that the first four moments of r_t, conditional to $\mathcal{F}_{t-1}$, can be calculated in closed form. The propositions introduced hereafter illustrate the resulting expressions. For simplicity of notation, the operators presenting a $t-1$ subscript correspond hereafter to conditional operators on $\mathcal{F}_{t-1}$, e.g., for a random variable X, $\mathbb{E}_{t-1}[X] = \mathbb{E}[X|\mathcal{F}_{t-1}]$.

2.3.1 Conditional expectation of return

First, we focus on the conditional expected value of return implied by our model.

Proposition 2 *The conditional expectation of return r_t is equal to*

$$\mu_t = \mathbb{E}_{t-1}[r_t] = \mu + \sigma\sqrt{\tau_t} \exp\left\{\frac{\gamma_1}{2} v_t^e\right\} \times \frac{1}{2}\rho\gamma_2\sigma_u \exp\left\{\frac{1}{8}\gamma_2^2\sigma_u^2\right\}, \tag{2.16}$$

where $\frac{1}{2}\rho\gamma_2\sigma_u \exp\left\{\frac{1}{8}\gamma_2^2\sigma_u^2\right\} = \mathbb{E}\left[\exp\left\{\frac{\gamma_2 v_t^u}{2}\right\} z_t\right]$ is the expectation of the NLN mixture.

Proposition 2 implies important features regarding the conditional mean of returns. First, the proposition highlights a natural risk-return relation, whose sign will depend on the parameters ρ and γ_2 (other components of the second term in the right hand side of equation 2.16 are all positive by definition). The relation between the conditional mean and the conditional volatility will be stronger (either positive or negative) as 1) the magnitude of the correlation between unexpected return and unexpected volume $|\rho|$ increases, 2) the impact of unexpected volume $|\gamma_2|$ increases, and 3) the standard deviation of volume innovations σ_u increases. Secondly, the conditional expectation of returns fluctuates over time according to the level of the expected volume of trading, and to τ_t the intensity of the variance of price increments. If v_t^e is constant over time, then τ_t drives the evolution of the conditional mean. A positive (resp. negative) risk-return relation implies that the higher τ_t the higher (resp. the lower) the expected return. On the other hand, if v_t^e also fluctuates over time, then the evolution of the conditional mean integrates the impact of the interaction with the expected volume of trading.

2.3.2 Higher moments

The following proposition derives the general expressions for higher moments of the return distribution.

Proposition 3 *For $n = 2, 3, 4$, the n^{th}-conditional centered moment of return r_t is equal to*

$$\mathbb{E}_{t-1}\left[(r_t - \mathbb{E}_{t-1}[r_t])^n\right] = \sigma^n \tau_t^{\frac{n}{2}} \exp\left\{\frac{n}{2}\gamma_1 v_t^e\right\} \times m_{NLN}^n$$

where $m_{NLN}^n = \mathbb{E}\left[\left(\exp\left\{\frac{\gamma_2 v_t^u}{2}\right\} z_t - \mathbb{E}\left[\exp\left\{\frac{\gamma_2 v_t^u}{2}\right\} z_t\right]\right)^n\right]$ is the n^{th} centered moment of the NLN mixture.

Proof

$$\mathbb{E}_{t-1}\left[(r_t - \mathbb{E}_{t-1}[r_t])^n\right] = \mathbb{E}_{t-1}\left[\sigma^n \tau_t^{\frac{n}{2}} \exp\left\{\frac{n}{2}\gamma_1 v_t^e\right\}\left(\exp\left\{\frac{\gamma_2 v_t^u}{2}\right\} z_t - \mathbb{E}\left[\exp\left\{\frac{\gamma_2 v_t^u}{2}\right\} z_t\right]\right)^n\right]$$
$$= \sigma^n \tau_t^{\frac{n}{2}} \exp\left\{\frac{n}{2}\gamma_1 v_t^e\right\} \mathbb{E}\left[\left(\exp\left\{\frac{\gamma_2 v_t^u}{2}\right\} z_t - \mathbb{E}\left[\exp\left\{\frac{\gamma_2 v_t^u}{2}\right\} z_t\right]\right)^n\right].$$

Given the analytic expressions of m_{NLN}^n for $n = 2, 3, 4$, the following corollary follows directly from Proposition 3, and allows to calculate conditional variance, skewness and kurtosis of returns.

Corollary 1 *The conditional moment coefficients of returns are given by the following expressions.*

- *The conditional variance of return* $\mathrm{Var}_{t-1}(r_t)$ *is equal to:*

$$\mathrm{Var}_{t-1}(r_t) = \sigma^2 \tau_t \exp\{\gamma_1 v_t^e\} \times m_{NLN}^2,$$

  where $m_{NLN}^2 = e^{\frac{1}{2}\gamma_2^2\sigma_u^2}\left[1 + \rho^2\gamma_2^2\sigma_u^2\left(1 - \frac{1}{4}\exp\left\{-\frac{1}{4}\gamma_2^2\sigma_u^2\right\}\right)\right]$.

- *The conditional skewness of return* $\mathrm{Skew}_{t-1}(r_t)$ *is equal to:*

$$\mathrm{Skew}_{t-1}(r_t) = \left(\frac{\sigma^2\tau_t \exp\{\gamma_1 v_t^e\}}{\mathrm{Var}_{t-1}(r_t)}\right)^{\frac{3}{2}} \times m_{NLN}^3 = \left(m_{NLN}^2\right)^{-\frac{3}{2}} \times m_{NLN}^3$$

  where

$$m_{NLN}^3 = \rho\gamma_2\sigma_u e^{\frac{9}{8}\gamma_2^2\sigma_u^2}\left[\left(1-\rho^2\right)A + \rho^2 B\right],$$
$$A = \frac{9}{2} - \frac{3}{2}e^{-\frac{1}{2}\gamma_2^2\sigma_u^2},$$
$$B = \left(\frac{9}{2} + \frac{27}{8}\gamma_2^2\sigma_u^2\right) - \frac{3}{2}\left(1+\gamma_2^2\sigma_u^2\right)e^{-\frac{1}{2}\gamma_2^2\sigma_u^2} + \frac{1}{4}\gamma_2^2\sigma_u^2 e^{-\frac{3}{4}\gamma_2^2\sigma_u^2}.$$

- *The conditional kurtosis of return* $\mathrm{Kurt}_{t-1}(r_t)$ *is equal to:*

$$\mathrm{Kurt}_{t-1}(r_t) = \left(\frac{\sigma^2\tau_t \exp\{\gamma_1 v_t^e\}}{\mathrm{Var}_{t-1}(r_t)}\right)^{2} \times m_{NLN}^4 = \left(m_{NLN}^2\right)^{-2} \times m_{NLN}^4$$

  where

$$m_{NLN}^4 = e^{2\gamma_2^2\sigma_u^2}\left[3\left(1-\rho^2\right)^2 + 6\rho^2(1-\rho^2)C + \rho^4 D\right],$$
$$C = (1 + 4\gamma_2^2\sigma_u^2) - \frac{3}{2}\gamma_2^2\sigma_u^2 e^{-\frac{3}{4}\gamma_2^2\sigma_u^2} + \frac{1}{4}\gamma_2^2\sigma_u^2 e^{-\frac{5}{4}\gamma_2^2\sigma_u^2},$$
$$D = \left(3 + 24\gamma_2^2\sigma_u^2 + 16\gamma_2^4\sigma_u^4\right) - \left(9 + \frac{27}{4}\gamma_2^2\sigma_u^2\right)e^{-\frac{3}{4}\gamma_2^2\sigma_u^2}$$
$$+ \frac{3}{2}\left(1+\gamma_2^2\sigma_u^2\right)\gamma_2^2\sigma_u^2 e^{-\frac{5}{4}\gamma_2^2\sigma_u^2} - \frac{3}{16}\gamma_2^4\sigma_u^4 e^{-\frac{3}{2}\gamma_2^2\sigma_u^2}.$$

First, note that the conditional variance implied by the model is time-varying whereas conditional skewness and conditional kurtosis are constant over time and therefore equal by definition to the unconditional skewness and kurtosis. In other words, while the expected return and the expected volatility are time-varying, the model preserves constant higher moments. Regarding the second moment, the two components of the variance affect the prediction. On the one hand, τ_t, which is completely determined given lagged information, impacts directly the variance forecast according to past information concerning price increments unrelated to volume. On the other hand, the volume component, split into a predictable part and an innovation part, makes conditional variance fluctuating according to future expected volume of trading and average squared normal log-normal innovation (i.e., second moment of the NLN mixture).

Interestingly, the parameters related to the unexpected part of g_t, i.e., γ_2, ρ, σ_u, clearly impact the third and fourth moments. Assuming σ_u is non-zero, we can see that the skewness and (excess) kurtosis are non-zero if γ_2 and ρ are non-zero. Furthermore, if $\gamma_2 > 0$ (as highlighted later in section 3.7), the sign of ρ determines the sign of the return's skewness. Its magnitude also depends on γ_2 and σ_u: the greater γ_2 and σ_u, the greater the absolute value of the skewness. We illustrate this result in Figure 2.1. The latter presents the analytic values of skewness and kurtosis of returns as functions of γ_2, ρ and σ_u. For each panel, except the varying parameter, all other parameters are kept constant and set to an arbitrary plausible value according to our empirical results over all stocks in our sample from 1994-01-03 to 2014-12-31[18], discussed latter in section 3.7. We can clearly observe that the magnitude of the skewness increases as the magnitude of γ_2, ρ and σ_u increases. The sign is determined according to the sign of ρ. As shown in panels B and F, the (excess) kurtosis is an increasing function in γ_2 and σ_u. However, depending on the value of γ_2, the kurtosis is not monotonic in ρ (cf., panel D).

Finally, as mentioned before and highlighted by the expressions of the conditional moments of return[19], if $\gamma_2 = 0$ and/or $\sigma_u = 0$, i.e., no impact of unexpected volume and/or no shocks in volume, the asset return is conditionally Gaussian, with neither asymmetry nor fat tails.

2.3.3 Covariance between return and volume

The initial motivation of MDH models is to develop a framework able to create correlation between volume and return's variance, or in other words correlation between volume and squared return. Using our parametric modeling we derive below the conditional covariance between return and volume as well as the covariance between squared return and volume.

Proposition 4 *The conditional covariance between return and volume is equal to*

$$\mathrm{Cov}_{t-1}(r_t, V_t) = \sigma\sqrt{\tau_t} e^{\frac{1}{2}(\gamma_1+2)v_t^e} V_{t-1} \times \ \mathrm{Cov}_{t-1}(e^{\frac{1}{2}\gamma_2 v_t^u} z_t, e^{v_t^u})$$

[18] Except in panels E and F, σ_u is set to its median values over all stocks: 0.35.

[19] The unconditional moments of returns can also be derived. The terms $\mathbb{E}\left[\sqrt{\tau_t}\right]$ and $\mathbb{E}\left[\tau_t^{\frac{n}{2}}\right]$ can be approximated numerically using for instance the expectation of a Taylor expansion of $\tau_t^{\frac{n}{2}}$ around $(\mathbb{E}[\tau_t])^{\frac{n}{2}}$.

where $\text{Cov}_{t-1}(e^{\frac{1}{2}\gamma_2 v_t^u} z_t, e^{v_t^u}) = \frac{1}{2}\rho\sigma_u e^{\frac{1}{8}(\gamma_2+2)^2\sigma_u^2}\left[2+\gamma_2\left(1-e^{-\frac{1}{2}\gamma_2\sigma_u^2}\right)\right]$.

This expression can be derived using the expression of the moments derived in section 2.3.2 and the fact that $V_t = V_{t-1}e^{v_t^e+v_t^u}$.

Two important features of the model are put forward by Proposition 4. First, given the positive sign of most terms, we can observe that the sign of the covariance between return and volume only depends on the sign of ρ and γ_2. As section 3.7 will show, γ_2 being positive for all stocks, the sign of the covariance will therefore depend on the correlation ρ. This result is quite intuitive as the latter parameter captures the correlation between shocks in return and shocks in volume. Secondly, the magnitude of the covariance is related to the different terms of the volatility of returns as well as past volume V_{t-1}. Hence, during periods characterized by a persistent high level of volatility, i.e., large τ_t, the relation between return and volume will be larger in magnitude. In addition, during periods of large volume (large V_{t-1}) the relation will also be more important.

The next proposition derives the expression of the conditional covariance between squared return and volume, i.e., the covariance between return's variance and volume.

Proposition 5 *The conditional covariance between squared return and volume is equal to*

$$\text{Cov}_{t-1}(r_t^2, V_t) = \sigma^2\tau_t e^{(\gamma_1+1)v_t^e} V_{t-1} \times \text{Cov}_{t-1}(e^{\gamma_2 v_t^u} z_t^2, e^{v_t^u}) \quad + \quad 2\mu\text{Cov}_{t-1}(r_t, V_t).$$

The first covariance term in Proposition 5 can also be derived using the expression of the moments in section 2.3.2 and the fact that $V_t = V_{t-1}e^{v_t^e+v_t^u}$.

Note that the first term in the right hand side of Proposition 5 cannot be negative, capturing the idea that positive shocks in squared return are associated with positive shocks in volume. The second term will depend on the sign of ρ (as the expression given in Proposition 4) and μ, and thus can amplify of reduce the covariance. As mentioned above for Proposition 4, the magnitude of the relation between squared returns and volumes is also related to the different components of the return's variance and volume.

2.4 Empirical analysis

This section presents the empirical analysis of the model, using U.S. large capitalization stocks. First, the empirical properties of data are discussed. Then, the Maximum Likelihood Estimation methodology of the model is presented, followed by a model selection procedure conducted among nested models of the specification previously introduced and the VA-GARCH of Lamoureux and Lastrapes (1990a). Finally, we derive conclusions regarding stock variance dynamics and its relation with trading volume according to fitted parameters.

2.4.1 Data

Data used in this section correspond to daily total (log) returns and trading volume (in millions of traded shares) of individual component stocks of the S&P 100 Index as of 2014-12-31, corrected for dividends and stock splits. Our sample begins on 1994-01-03 and ends on 2014-12-31 (twenty years). Stocks with historical data sample starting after 1994-01-03 are removed. The list of the remaining 80 stocks studied in this section is provided in Table 2.1. For space consideration in our Tables, we provide detailed statistics related to the twenty stocks with highest mean daily volume over the period 1994-2014, as well as summary statistics over all stocks considered[20].

Table 2.2 reports sample statistics of return time series. The sample standard deviation averages to 2.15% over all series. Most stocks display a negative skewness (50 out of 80). This coefficient ranges from -4.86 for RTN to 3.93 for MO. The (excess) kurtosis is positive for all assets. Finally, except TWX, the return time series are all stationary according to KPSS statistical test[21].

In the left part of Table 2.3 we present sample statistics of difference in log-trading volume. The skewness of the series is mainly positive (76 stocks out of 80) and is close to zero (between -0.14 for BRK and 0.60 for AIG), while the (excess) kurtosis is slightly positive for all assets (unreported for space consideration). All changes in log-volume are stationary according to KPSS test.

2.4.2 Maximum Likelihood Estimation

We present hereafter the Maximum Likelihood Estimation (MLE) approach used to estimate the model. In order to avoid parameter identification issues due to the products in equation (2.11), and given that elements of the volume's ARMA structure are embedded in the return dynamics, we use a two-step estimation procedure (cf., Heckman, 1979, for instance). First, we estimate the ARMA models described by equation (2.9), by maximizing the log likelihood:

$$\log L_1(\phi, \theta, \sigma_u) = \sum_{t=2}^{T} \log \text{pdf}_{\log V_t|\mathcal{F}_{t-1}}(\log V_t|\phi, \theta, \sigma_u), \tag{2.17}$$

where T stands for the sample size, and ϕ, θ respectively for the AR and MA parameters. Note that $\text{pdf}_{\log V_t|\mathcal{F}_{t-1}}$ corresponds to a Gaussian density function.

In a second step, using the fact that the distribution of $r_t|\mathcal{F}_{t-1}, v_t^u$ is Gaussian, we estimate the remaining parameters $\Theta = [\mu, \sigma, \alpha_1, \alpha_2, \beta_1, \gamma_1, \gamma_2, \rho]'$ maximizing the log likelihood:

$$\log L_2(\Theta, \hat{\phi}, \hat{\theta}, \hat{\sigma_u}) = \sum_{t=2}^{T} \log \text{pdf}_{r_t|\mathcal{F}_{t-1}, v_t^u}(r_t|\Theta, \hat{\phi}, \hat{\theta}, \hat{\sigma_u}), \tag{2.18}$$

[20] Detailed numerical results concerning all stocks are available upon request.
[21] See Kwiatkowski, Phillips, Schmidt, and Shin (1992) on the KPSS test.

where

$$\mathrm{pdf}_{r_t|\mathcal{F}_{t-1},v_t^u}(x|\Theta,\phi,\theta,\sigma_u) = \frac{1}{\sqrt{2\pi\left(1-\rho^2\right)\sigma^2\tau_t g_t}} \exp\left\{-\frac{1}{2}\frac{(x-\mu_t-\frac{\rho}{\sigma_u}\sigma\sqrt{\tau_t g_t}\times v_t^u)^2}{\left(1-\rho^2\right)\sigma^2\tau_t g_t}\right\}.$$

As the standard regularity conditions are met for both log-likelihood functions L_1 and L_2, then the second-step maximum likelihood estimator is consistent and asymptotically normally distributed (see Murphy and Topel, 2002).

The selected ARMA models, according to L_1 and AIC criterion, are presented in the right part of Table 2.3. Most stock volumes imply an autoregressive structure with two lags, and a moving-average with three lags. We note that the AR(1) coefficient is positive for the vast majority of series (more than 90%), with an average value of 0.56, exhibiting a moderate but non-negligible level of persistence in the variations of log-volume. The AR(2) coefficient (when selected) are also mainly positive. Interestingly, the MA(1) coefficient is large and negative for all assets (except BAC), with an average value of -1.12, meaning that a large positive (resp. negative) innovation will tend to decrease (resp. increase) the predicted volume for the next period. In other words, if the unexpected shock in the volume was large (positive) yesterday, we expect a lower volume today, and vice versa. The coefficients corresponding to further moving average lags are much lower in magnitude and positive on average.

We note that the fitted values of σ_u are all significant and quite concentrated between 0.3 and 0.4 (average value of 0.36 over all stocks), illustrating the fact that the range of unexpected volume of trading is quite similar across this sample of highly liquid stocks.

Finally, the R^2 of the ARMA model, defined as:

$$R^2 = 1 - \frac{\mathrm{Var}(u_t)}{\mathrm{Var}(\log V_t/V_{t-1})},$$

are also reported in the last column of Table 2.3. The fraction of variations in volume captured by the ARMA structure is about one fourth (0.26). Therefore, according to our specification, a large part of log volume is due to innovations. However given the level of R^2 the decomposition of volume between a predictable and an unexpected component appears to be worthy. Indeed, ignoring the splitting would imply presumably to capture the impact of volume through the innovation and not the expected component, which is a priori also important. Our estimations (presented later) support clearly the fact that expected and unexpected volume do not have a similar impact on volatility.

2.4.3 Model selection procedure

Looking at equations (2.8) and (2.15), the model has numerous variants depending upon the restrictions imposed on the different parameters, especially on the impact of expected and unexpected volume. In order to investigate the different alternatives to implement the model and avoid overfitting, we proceed below to a model selection procedure among nested models. In addition, as a benchmark model combining volume and GARCH speci-

fication, we also estimate the VA-GARCH model[22] of Lamoureux and Lastrapes (1990a).

For each stock, using the MLE procedure previously described, the following competing models are fitted[23] 1) unrestricted VF-GARCH, 2) VF-GARCH with $\gamma_2 = 0$ (only expected volume impact), 3) VF-GARCH with $\gamma_1 = 0$ (only unexpected volume impact), 4) VF-GARCH with $\gamma_1 = \gamma_2$ (same impact of expected and unexpected volume), 5) GJR-GARCH, i.e., $\gamma_1 = \gamma_2 = \rho = 0$, 6) VF-GARCH without GARCH component ($\alpha_1 = \alpha_2 = \beta_1 = 0$) and 7) the VA-GARCH of Lamoureux and Lastrapes (1990a). Using the likelihood estimates[24], the different specifications are compared using Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). For a given model, these statistics are calculated as follows:

$$AIC = 2K - 2\log L_2(\hat{\Theta}, \hat{\phi}, \hat{\theta}, \hat{\sigma_u}),$$
$$BIC = K\log(T) - 2\log L_2(\hat{\Theta}, \hat{\phi}, \hat{\theta}, \hat{\sigma_u}),$$

where K represents the number of parameters of the specification, and $\hat{\Theta}$ the fitted parameters set for returns. In addition, Likelihood Ratio Tests (LRT) (at 5% level) are performed. For two competing nested models, twice the difference in log-likelihoods is compared to the critical value of a χ^2 distribution with degree of freedom equal to the difference in the number of parameters between the two models. Table 2.4 summarizes the results. Each panel corresponds to a given selection criterion. In a given panel, each number corresponds to the percentage of stocks for which the model specified as row name is selected with respect to the model specified as column name[25].

The unrestricted VF-GARCH model appears to be the most satisfactory specification. Indeed, the three selection criteria favor this specification for all assets with respect to the VF-GARCH with $\gamma_2 = 0$, the VF-GARCH with $\gamma_1 = \gamma_2 = \rho = 0$, the VF-GARCH with $\alpha_1 = \alpha_2 = \beta_1 = 0$, and the VA-GARCH. Furthermore, for most stocks, the unrestricted specification is also selected with respect to the other variants. Note that the specification $\gamma_1 = 0$, implying only an impact of unexpected volume, is largely selected relative to other models (except the unrestricted one). These observations already suggest that both expected and unexpected volume of trading are important factors in order to determine stock variance, but contemporaneous innovation in volume seems to have a greater importance than the predictable part of this variable. In addition, as suggested by the results of the VA-GARCH with respect to the VF-GARCH specifications, the decomposition of the variance between a persistent GARCH and a volume component (instead of adding a volume component in the GARCH variance equation) is clearly motivated by the data.

According to the results presented above, the next section discusses fitted parameters and variance estimates of the unrestricted VF-GARCH model, as well as its implications

[22]We use trading volumes in billions of traded shares.

[23]Given our two-step procedure, we do not have identification problems under the different null hypothesis on the parameters, cf., Hansen (1996) on this issue.

[24]The likelihood estimates of the different VF-GARCH models are adjusted for the term related to v_t^u for comparison with the VA-GARCH.

[25]The LRT panel is less complete than others as Likelihood ratio tests can only be performed with nested models.

on the relationship between volatility and volume.

2.4.4 Fitted parameters and conditional estimates

This section presents our empirical findings using the VF-GARCH model. We estimate the parameters of the unrestricted model using numerical techniques to maximize the likelihood functions for each stock over the full period 1994-2014. Sub-samples are considered in section 2.5 in order to derive empirical regularities across the different cycles in our sample.

In Table 2.5 we present the VF-GARCH fitted parameters over the period 1994-2014. t-tests statistical significance at 1%, 5% and 10% are respectively indicated by ***,**, and *. We discuss the results in the next paragraphs.

Volume component parameters

First, the fitted parameters regarding g_t the volume component of the variance are analyzed. Clearly, impacts of both expected and unexpected volume of trading appear to be substantial for all stocks of the S&P 100, as suggested by the percentage of stocks for which γ_1 and γ_2 are significant (100%). Interestingly, the signs of these two parameters are completely different. While γ_1 is negative for all stocks, γ_2 is always positive. This observation is confirmed and highlighted by Figure 2.2. The latter presents a scatter plot of fitted values of γ_1 and γ_2 for all stocks in the sample over the period 1994-2014. We can even discern a relationship between the magnitude of the two parameters: the lower the impact of expected volume, the lower the impact of unexpected shocks.

Given the expression of g_t (equation 2.8), the first conclusion is that expected and unexpected volume have completely different impacts on variance: 1) a high expected volume tends to decrease volatility, while 2) a high unexpected volume increases volatility. In other words, if a large number of trades is anticipated, the asset volatility is reduced. On the other hand, if a large number of unanticipated trades occurs, this has the effect of increasing volatility. These two opposite effects are not contradictory with the widespread conclusion of previous studies (see for instance Gallant et al., 1992; Harris and Raviv, 1993; Shalen, 1993) that high volume tends to increase volatility. Indeed, as shown by the ARMA structures presented in Table 2.3, for all stocks the expected volume is greatly impacted by the lagged innovation, i.e., the lagged unexpected volume, through the negative MA(1) coefficient. Hence, if we observed yesterday a large unexpected shock in volume, we anticipate a large decrease in expected volume and therefore an increase in expected volatility[26]. This observation underlines the importance of the decomposition of volume and puts forward the fact that volatility will be presumably higher during periods of large (positive) innovations in volume. This reveals the usual conclusion that large price movements are accompanied by large shocks in volumes. Note that a non-zero value of γ_2 is also associated with a positive conditional excess kurtosis as highlighted in Corollary

[26]On the other hand, if the last past shock in volume was unusually low, i.e., large negative unexpected volume, the expected volume for the next day will be presumably high due to the negative MA(1) coefficient, implying a decrease in the expected volatility.

1. Hence the model indicates that the conditional distributions of all assets in the sample are leptokurtic.

It is also important to focus on the fitted values of ρ, which are significant for most assets (52%). Indeed, as γ_2 is obviously positive across stocks, a direct conclusion according to section 2.3 is that the correlation between unexpected return and unexpected volume is the key driver of the sign of the risk-return relation (cf., Proposition 2), and also determines the sign of the conditional skewness of returns (cf., Corollary 1). Table 2.5 suggests that ρ is predominantly slightly positive in our full sample, with an average value of 0.023 over all stocks. Except for GE and MO, all significant fitted values of ρ are positive. The conclusion is that the correlation between unexpected volume and unexpected return is predominantly positive implying a positive risk-return trade-off at the daily horizon, for most assets on this sample. Note that, if one assumes that shocks in volume are a good proxy for news shocks then this positive correlation indicates that on average stock price changes are positively affected by new information. This observation can be related to the model of Campbell et al. (1993), implying that *"a stock price decline on a high-volume day is more likely than a stock price decline on a low-volume day to be associated with an increase in the expected stock return"* (Campbell et al., 1993, p. 905). We use Proposition 2 and we assume that a negative shock in the lagged return implies a greater increase in τ_t the GARCH component than a positive shock, i.e., leverage effect (see later). First, it is clear that if $\rho > 0$, a greater value of τ_t due to a last large negative return implies a greater expected stock return. In addition, if the last shock in volume was large and positive, the expected volume will decrease (cf., above), implying a greater value of the term $\exp\{\frac{\gamma_1}{2} v_t^e\}$. As $\rho > 0$, this will therefore generate an even greater increase in the expected return. Hence similarly to Campbell et al. (1993), we find that a stock price decline on a *high-unexpected volume day* is more likely than a stock price decline on a *low-unexpected volume day* to be associated with an increase in the expected stock return (implying potentially return reversals).

GARCH component parameters

The fitted parameters of the GARCH component, i.e., $\alpha_1, \alpha_2, \beta_1$, are now examined. As a benchmark model without volume, fitted parameters of the GJR-GARCH, i.e., restricted VF-GARCH model with $\gamma_1 = \gamma_2 = \rho = 0$, are presented in Table 2.6 using the same sample. This model exhibits strong ARCH effect, i.e., positive autocorrelation in squared returns. Indeed, the average values of α_1 and β_1 are respectively 0.035 and 0.923, with significance rates of 96% and 100% over all stocks. This result confirms that the persistence of variance is a stylized feature of stock returns. Moreover, the parameter α_2, capturing the asymmetric effect in volatility (a.k.a., leverage effect), is statistically significant for 99% of stocks, with an average value of 0.0623. This indicates that, under a GARCH-type specification, lagged negative innovations in return tend to have a greater impact on assets' variance than positive lagged innovations.

Focusing now on Table 2.5 corresponding to the VF-GARCH model, the same parameters are inspected. Interestingly, the percentages of stocks for which α_1 and β_1 are

significant are quite similar to those discussed above: 95% for α_1 and 100% for β_1. In addition, the average values of these parameters are not dramatically reduced compared to the GJR-GARCH. The mean of α_1 is even higher: 0.0815, while the mean of β_1 is 0.87. These results suggest that ARCH effect is not reduced when accounting for the impact of trading volume, and that variance persistence is captured through the GARCH component of our model[27]. Therefore, a large fraction of the persistence of volatility is unrelated to trading volume and this behavior remains, whether or not volumes are included in the specification. This important result is in line with recent studies such as Fleming et al. (2006), but is in contradiction with authors arguing that volume subsumes ARCH effect (see Lamoureux and Lastrapes, 1990a).

Figure 2.3 displays the persistent part of the expected volatility of the VF-GARCH, i.e., $\sigma\sqrt{\tau_t}$ (dark line) for APPL and F stocks. This component behaves relatively similarly to the GJR-GARCH estimate (grey line), but interestingly, it appears to be even smoother. Several large spikes exhibited by the GJR-GARCH model seem to be reduced or cleared by the filtered component of the VF-GARCH model, especially in September 2000 for APPL[28]. According to the latter model, this implies that these extreme increases in predicted volatility are not associated with the persistent component of volatility unrelated to trading volume. Note that the mean level of $\sigma\sqrt{\tau_t}$ is slightly lower than the average value of the GJR-GARCH volatility. Indeed, in the "total" variance prediction of the VF-GARCH, the filtered component will be amplified by both the expected volume and the second moment of the normal log-normal mixture.

Finally, when looking at parameter α_2 in Table 2.5, conclusions regarding the impact of trading volume on the usual asymmetric effect in volatility can be discussed. Indeed, the fitted values of α_2 are relatively close between the GJR-GARCH and the VF-GARCH, with an average value of 0.073 for the latter, and a significance rate of 86%. One cannot conclude that the integration of volume in the variance dynamic substantially attenuates the asymmetric effect in volatility as documented in Gallant et al. (1992). The usual leverage effect appears to be a feature of stock variance unrelated to trading volume.

Contribution of trading volume

We focus now on the contribution of the different components, i.e., τ_t and g_t, to the total variance. In other words, how much of stock variance can be explained by expected and unexpected volume, and how much is due to variations unrelated to trading activity and information arrival. To investigate this question, following Engle et al. (2013), we calculate empirically for each asset the following variance ratios using the fitted parameters and

[27]We do not find that the fitted values of β_1 in the VA-GARCH model of Lamoureux and Lastrapes (1990a) are dramatically lower. The average value of α_1 and β_1 are respectively 0.17 and 0.71.

[28]APPL market share value fell by 51.9% on September 29, 2000.

volatility:

$$\text{Var}(\log(\tau_t))/\text{Var}(\log(\tau_t g_t)), \tag{2.19}$$

$$\text{Var}(\log(g_t))/\text{Var}(\log(\tau_t g_t)), \tag{2.20}$$

$$\text{Var}(\gamma_1 v_t^e)/\text{Var}(\log(\tau_t g_t)), \tag{2.21}$$

$$\text{Var}(\gamma_2 v_t^u)/\text{Var}(\log(\tau_t g_t)), \tag{2.22}$$

as respective contributions of the persistent GARCH component, volume component (both expected and unexpected), expected volume, and unexpected volume component. Note that the variations explained by τ_t can be considered as unrelated to trading volume. The values of these ratios are summarized in Table 2.7.

The estimates tell us that both the GARCH factor and the volume component are key contributors to the total variance for the stocks in the sample. On average g_t contributes to 36% of variations in asset variance. However, the volume contribution is greater than 50% for only eight stocks. Therefore, the trading activity is an important determinant of the stock variance (about one third), but the latter is also clearly driven by the persistent component τ_t unrelated to volume. Note that the contribution of τ_t seems to be very high for the stocks of banks in our sample (75% for BAC, 85% for WFC, 70% for JPM, for instance), highlighting that the variability of these assets are highly affected by factors unrelated to volume.

Furthermore, focusing on the contribution of the different parts of volume, one can see that the variations of the unexpected volume component are substantial in order to understand the variance dynamics. Indeed, the unexpected volume component has an average contribution of 32% across stocks, this component appears to clearly drive the volume component. This result is consistent with Fleming et al. (2006)'s finding that "*the unpredictable component of volatility is closely tied to the unpredictable component of contemporaneous trading volume*" (Fleming et al., 2006, p. 1589). Interestingly, the contribution of the expected volume component is much lower.

Hence two main conclusions arise out of this analysis, for most stocks: 1) both components (persistent GARCH and volume) are important drivers of volatility dynamics, and 2) the majority of variations in the volume component can be attributed to unexpected volume of trading. The first conclusion implies that studying only volume is insufficient in order to understand stock market volatility dynamics and highlights the limits of standard MDH models, while the second conclusion implies that a large portion of daily variance is a priori unpredictable as it is due to innovations in volume, independent of the lagged available information.

In order to illustrate these results, Figure 2.4 displays the different components of volatility for BAC and CAT time series. As expected, the VF-GARCH total volatility is much less smooth than the filtered component $\sigma\sqrt{\tau_t}$. In fact, many transient spikes exhibited by the total volatility, are not present in the dynamic of τ_t. The intuitive interpretation is that these large variations in conditional stock variance can be attributed to trading volume and not to shocks in the persistent mean-reverting volatility of price

increments. Basically, assuming a large negative return is observed at time $t-1$, if this large negative return was accompanied by a large (positive) unexpected volume of trading, the component τ_t will not be considerably impacted by the return's shock, given the scaling in the dynamic of τ_t (cf., equation 2.15). Hence, the grey spikes observed in Figure 2.4 are due to the volume component but do not persist given the small weight of innovations beyond lag 1 in the ARMA structure of (log) volumes (cf., Table 2.3). Therefore, according to the VF-GARCH model, the large spikes observed in the expected volatility time series are much more transient than suggested by a GARCH-type model, agnostic on the relationship between trading volume and stock variance.

Clearly, the variations of the dark line (due to τ_t) exhibit much smaller amplitudes for CAT than for BAC stock, as expected given the small contribution of volume for the latter. Figure 2.4 also gives an idea of the level of the different variance components across time. In fact, one can see that during the financial crisis period 2007-2008, the level of τ_t was much more above its mean for BAC than for CAT stock.

2.5 Structural breaks and sub-samples

Note that, our initial sample (1994-2014) contains recent data but also presents several business cycles of economic expansions and recessions, raising concerns about potential breaks. Structural breaks in volatility have been highlighted by several studies: Lamoureux and Lastrapes (1990b) and Andreou and Ghysels (2002) for instance. In particular, the former paper shows that structural breaks in the unconditional level of GARCH variance can lead to over-estimation of the persistence in GARCH(1,1) model over long periods. Therefore, we consider several sub-samples in order to study breaks and empirical regularities.

2.5.1 Structural breaks in volatility

We split our 20-years sample between five periods: three periods of economic expansions (January 1994 to February 2001, December 2001 to November 2007 and July 2009 to December 2014) and two periods of economic recessions[29] (Mars 2001 to November 2001 and December 2007 to June 2009). Then, for all stocks in our initial sample, we maximize the log-likelihood functions previously introduced using the different sub-samples. In order to test for structural breaks, we compute the following likelihood ratio statistic, comparing the log-likelihood of the full sample with those of the sub-samples:

$$2\left[\ell(\hat{\Theta}_{full}) - \sum_{i=1}^{5} \ell(\hat{\Theta}_i)\right] \sim \chi^2_{K\times 4}, \tag{2.23}$$

where $\ell(\hat{\Theta}_{full})$ is the log-likelihood of the full-sample, $\ell(\hat{\Theta}_i)$ the log-likelihood on the sub-sample i, and K the number of parameters in the model.

[29]The dates of the business cycles are borrowed from NBER.

The results of the analysis show that a vast majority of stocks (79%) rejects the absence of structural breaks. This result is not completely unexpected as the specification of τ_t is similar to a GARCH model, implying the same potential sources of breaks. Indeed we find that the GJR-GARCH exhibits breaks for 78% of assets. Due to this result, we also focus our empirical analysis on the fitted parameters using the sub-samples in order to verify our previous findings. Note that when performing the model selection procedure described in section 2.4.3 over each sub-period, we find that the unrestricted VF-GARCH model is always the best choice except during the recession period of 2001[30].

2.5.2 Fitted parameters

We focus first on the parameters interacting with trading volumes in g_t. Figure 2.5 presents the scatter-plots of fitted values of γ_1 and γ_2 for the different sub-periods of interest. Clearly our previous conclusions regarding the signs and the interpretation of the parameters remain valid. Indeed, for all sub-periods the significant fitted values of γ_1 are negative (except for respectively two, two, and one stock(s) in panels B, D and E), meaning that expected volume of trading are associated with a reduction of expected volatility. Note that most parameters appear to lie between 0 and -2. Regarding the second parameter γ_2, it is always positive and mainly above one. This confirms our previous finding that unexpected shocks in trading volume increase volatility.

We present a summary of the fitted parameters for the VF-GARCH model over the different sub-samples in Table 2.8. Again, as a benchmark model, we also present the fitted parameters of the restricted model with $\gamma_1 = \gamma_2 = \rho = 0$ (GJR-GARCH) over the same sub-samples in Table 2.9. Interestingly, we observe that the sign of (average) ρ is not always positive, depending on the sub-period. In particular, ρ is predominantly negative for the periods 2001-03 to 2001-11, 2007-12 to 2009-06, and 2009-07 to 2014-12. Assuming that shocks in volume are a good proxy for news shocks then this indicates that on average stock price changes are negatively impacted by new information. In other words, bad news are predominant during these periods. While this result appears intuitive for the two periods 2001-03 to 2001-11 and 2007-12 to 2009-06 as they correspond to recessions, it is more surprising for the sample 2009-07 to 2014-12.

Our previous conclusions on the component τ_t remain valid. First, the parameter β_1 is large and significant: above 0.8 and significant for at least 95% of stocks in all sub-samples, highlighting an important level of persistence in the GARCH component of the VF-GARCH model. This feature is also present in the model without volume, exhibiting similar values for β_1 in Table 2.9. Note that the lowest value of β_1 is obtained for the shortest period for both the GJR-GARCH and the VF-GARCH, as suggested by Lamoureux and Lastrapes (1990b). Regarding α_1, while its mean level is very similar across sub-samples (between 0.07 and 0.09), its significance rate across stocks oscillates between 34% during the recession of 2001 and 92% during the first sub-period 1994-2001. Similar observations can be drawn from the restricted model without volume. This further demonstrates the ability of the VF-GARCH model to account for persistence in volatility,

[30]The best model over this period is the VF-GARCH with $\gamma_1 = 0$.

captured by standard ARCH specifications.Finally, the mean value of α_2 also remains quite concentrated between 0.06 and 0.11 over the different periods, with significance rates between 51% and 75%. Note that the sum of the average fitted values $\alpha_1 + \beta_1 + \frac{1}{2}\alpha_2$ is very close to one and is respectively equal to 0.95, 0.93, 0.98, 0.96 and 0.96 over the five sub-samples. Therefore, we conclude that the persistence is a feature of volatility present in the different considered sub-periods, confirming that integrating trading volume in our model does not annihilate ARCH effect.

2.5.3 Volatility drivers

As described in section 2.4.4, Table 2.10 summarizes the contribution of the different components over the different sub-periods according to the variance ratios previously introduced. First, as previously discussed, a substantial fraction of variations in volatility appears to be unrelated to trading volume, underlying the importance of the variations in the component τ_t. However, we observe that unexpected shock in volume is also an important contributor to variance dynamics. Indeed, as before, the contribution of the unexpected volume component is much larger than the contribution of expected volume of trading. Interestingly, the average contribution of the volume component estimated over the different sub-samples is greater than the mean contribution over the full sample (36%, as presented in Table 2.7). This result is not surprising given the larger average fitted values of γ_2 for four sub-samples out of five (respectively 1.62, 1.83, 1.81 and 2.04 for the last four sub-samples), with respect to the mean value of γ_2 over the full sample (1.56). In fact, we can even see that for most stocks the volume component is the main driver of volatility dynamics (contribution greater than 50%) for all sub-periods except during the recession 2007-2009. The latter period is therefore characterized by important variations in the volatility component τ_t, underlying a high level of uncertainty, independent of the trading activity. Therefore, our previous conclusions remain: 1) a large fraction of volatility is a priori unpredictable as it is associated with contemporaneous innovations in trading volume, and 2) the variations in both the persistent component τ_t and the unexpected volume component remain substantial in order to explain volatility dynamics.

2.6 The effect of unexpected volume on volatility

This section discusses the impact on VF-GARCH volatility estimates of unexpected volume of trading occurring during important news events. Indeed, Andersen and Bollerslev (1998) and Maheu and McCurdy (2004) show that special announcements such as earning surprises can be associated with large but short-lived variations of volatility for exchange rates and equity respectively. In our model, if earning surprises are associated with large unexpected shocks in volume, our estimates of volatility will increase due to the volume component, but quickly reduce the following days.

To illustrate how special news events are captured by our model, we focus on three dates of earning surprises for IBM. We borrow from Maheu and McCurdy (2004) the

following dates of earning surprises[31]: 2000-07-20, 2000-10-18 and 2001-01-18. Table 2.11 presents the price and volatility variations as well as unexpected volumes during these periods. The exact dates of earning surprise are represented in bold. They are associated with large daily (absolute) returns: respectively 12.26%, -16.89% and 11.35%. Interestingly, we observe that the total variance of our model $\sigma^2\tau_t g_t$ also increases strongly on these specific dates. However, these variations can be largely attributed to the changes in g_t. The latter exhibits an increase of more than 250% for each specific date. In fact, we record large unexpected volume of trading on each specific announcement date. On the other hand, the day after the earning surprise, the component τ_t does not dramatically rise despite the large lagged shock, with respect to g_t. Indeed, as exhibited by equation (2.15) and Table 2.11, as the earning surprises are associated with large unexpected volume of trading (implying a large value of g_{t-1}), the impact of the return r_{t-1} in the dynamics of τ_t is limited.

We further illustrate this fact in Figure 2.6. The evolution of daily returns, unexpected volumes and volatility components are presented during the second earning surprise in October 2000 (grayed area). Clearly, the large negative return of 2000-10-18 (Panel A) is accompanied by a large unexpected volume (Panel B). The impact of the latter on the volatility is clear: the total volatility increases (due to g_t), while τ_t remains quite stable. Therefore, the total volatility quickly reduces the days after the surprise, and does not persist at a high level unlike estimates of the GJR-GARCH model presented in the Panel D. The Figure highlights the ability of the VF-GARCH model to disentangle the shocks in prices which are likely to be short-lived given the behavior of volumes.

Given equation (2.15), the impact of a return shock occurring at time t in the value of τ_{t+k} is given by $\beta_1^{k-1}(\alpha_1 + \alpha_2 \mathbb{1}_{r_t<\mu_t})\frac{(r_t-\mu_t)^2}{\sigma^2 g_t}$ for $k \geq 1$. Table 2.12 reports the numerical value of this term for $k = 10, 20, 30$ for the three studied earning surprises. While the impacts of the corresponding returns (in bold) appear to be quite large relative to other price variations after 10 days, most of the effect has vanished after 20 and 30 days.

The distinction of the model between long-lasting and short-lived shocks can also be observed on Figure 2.7. For the three largest negative returns occurring during the period 1994-01 to 2001-02, the Figure plots the rate of decay of the shock in τ_t component, i.e., the value of $\beta_1^{k-1}(\alpha_1 + \alpha_2 \mathbb{1}_{r_t<\mu_t})\frac{(r_t-\mu_t)^2}{\sigma^2 g_t}$ for $1 \leq k \leq 50$. Note that the two largest absolute returns (-16.89% and -16.2%) exhibit very different rate of decay. Moreover, the curves corresponding to the shocks -16.2% and -9.29% are very similar despite the difference in the magnitude of returns. When looking at unexpected volumes on these specific dates, we find that these three large shocks are associated with large innovations in volume: 1.36, 1.98 and 1.66 respectively for 2000-10-18, 1999-10-21 and 1996-04-17. In fact, the second and third realized unexpected volumes (1.98 and 1.66) are among the three largest unexpected volumes recorded over the period. Hence, the larger shock in volume for the second date justifies the much lower impact of the corresponding return with respect to the shock -16.82%.

[31] Maheu and McCurdy (2004) obtain earning surprises based on the average of analysts' forecasts reported by IBES.

Conclusion

The joint study of equity volatility and trading volume has been motivated by several episodes of large price variations accompanied by unusually high volume. This empirical association can be justified within a Mixture of Distribution Hypothesis (MDH) framework. Basically, these models predict that both volume and volatility are related to the unobservable news arrival rate. However, important stylized facts regarding stocks variance, notably persistence, are not captured by these specifications.

This paper introduces a parametric model, referred as Volume Filtered-GARCH (VF-GARCH), for studying the heteroskedasticity of individual stock returns, combining both MDH and GARCH models. Based on a variant of the Mixture of Distribution Hypothesis, the specification implies two components for the stock variance. The first one, modeled as an asymmetric GARCH, captures the persistence of the volatility of price variations unrelated to news arrival. On the other hand, the second component, the stochastic mixing variable, is modeled as a combination of both expected volume of trading and unexpected shocks in volume. These two quantities correspond respectively to the predictable part and the innovation of volume variations obtained after fitting an ARMA model to log-volume difference series.

The VF-GARCH specification implies an explicit association between volatility and volume, as well as the empirically motivated persistent mean-reverting behavior of stock variance. In addition, allowing shocks in the return and shocks in the volume to be correlated, the model leads to a normal log-normal mixture for the conditional distribution of returns, implying in particular a time-varying conditional mean and a non-zero skewness and (excess) kurtosis. According to the model, this reveals a natural risk-return relation capturing a result similar to Campbell et al. (1993)'s that *a stock price decline on a high-unexpected volume day is more likely than a stock price decline on a low-unexpected volume day to be associated with an increase of expected return.*

The main results of this paper concern the impact of trading volume on stock variance. We find that a large expected volume of trading decreases volatility, while large unexpected shocks in volume increase volatility. Given the structure of expected volume, this confirms that large unexpected shocks in volumes are accompanied by increase in stock variance.

Furthermore, we find that stylized features of volatility, that are ARCH effect and leverage effect persist even when volumes are incorporated into the model. In other words, the usual persistence of volatility and asymmetric impact of lagged return shocks are mainly associated with the component of variance unrelated to trading volume.

Empirical results show that the volume component is a key driver of stock variance. More specifically, the component of the model impacted by unexpected volume of trading appears to be the main contributor to variations in asset volatility in four business cycles out of five from 1994 to 2014. The latter result implies that a large portion of daily variance is a priori unpredictable as it is due to innovations in volume independent of lagged available information. On the other hand, the non-negligible remaining fraction of variations in stock variance is clearly associated with the persistent mean-reverting

component unrelated to trading volume, underlying the important role of the GARCH component in our model.

Finally, the model allows to analyze the variations of volatility during important news events. Indeed, using the example of IBM, we find that large price movements occurring during earning surprises are accompanied by high unexpected shocks in volume, implying large but short-lived increases in volatility, captured by the volume component of our model. The persistent GARCH component of the model remains stable during these periods.

Future works could exploit recent research on the relation between stock market volatility and macroeconomic variables in order to further disentangle the long-run movements of equity variance driven by macroeconomic fundamentals and the short-run movements driven by news arrival.

Tables and Figures

Nb.	Ticker	Name	Nb.	Ticker	Name
1	ORCL	ORACLE CORP	41	DIS	DISNEY WALT CO
2	MSFT	MICROSOFT CORP	42	BH	BIGLARI HOLDINGS INC
3	HON	HONEYWELL INTERNATIONAL INC	43	HPQ	HEWLETT PACKARD CO
4	EMC	E M C CORP MA	44	BAX	BAXTER INTERNATIONAL INC
5	KO	COCA COLA CO	45	OXY	OCCIDENTAL PETROLEUM CORP
6	DD	DU PONT E I DE NEMOURS & CO	46	WFC	WELLS FARGO & CO NEW
7	XOM	EXXON MOBIL CORP	47	APA	APACHE CORP
8	GD	GENERAL DYNAMICS CORP	48	MCD	MCDONALDS CORP
9	GE	GENERAL ELECTRIC CO	49	JPM	JPMORGAN CHASE & CO
10	IBM	INTERNATIONAL BUSINESS MACHS COR	50	UNP	UNION PACIFIC CORP
11	PEP	PEPSICO INC	51	TGT	TARGET CORP
12	MO	ALTRIA GROUP INC	52	BK	BANK OF NEW YORK MELLON CORP
13	COP	CONOCOPHILLIPS	53	LLY	LILLY ELI & CO
14	AMGN	AMGEN INC	54	WMT	WAL MART STORES INC
15	SLB	SCHLUMBERGER LTD	55	NKE	NIKE INC
16	CVX	CHEVRON CORP NEW	56	AXP	AMERICAN EXPRESS CO
17	AAPL	APPLE INC	57	INTC	INTEL CORP
18	TXN	TEXAS INSTRUMENTS INC	58	BAC	BANK OF AMERICA CORP
19	CVS	C V S HEALTH CORP	59	MDT	MEDTRONIC INC
20	BRK	BERKSHIRE HATHAWAY INC DEL	60	FDX	FEDEX CORP
21	UTX	UNITED TECHNOLOGIES CORP	61	LOW	LOWES COMPANIES INC
22	PG	PROCTER & GAMBLE CO	62	NSC	NORFOLK SOUTHERN CORP
23	SO	SOUTHERN CO	63	VZ	VERIZON COMMUNICATIONS INC
24	CAT	CATERPILLAR INC	64	T	A T & T INC
25	CL	COLGATE PALMOLIVE CO	65	USB	U S BANCORP DEL
26	BMY	BRISTOL MYERS SQUIBB CO	66	HD	HOME DEPOT INC
27	WBA	WALGREEN BOOTS ALLIANCE INC	67	AIG	AMERICAN INTERNATIONAL GROUP INC
28	BA	BOEING CO	68	MS	MORGAN STANLEY DEAN WITTER & CO
29	ABT	ABBOTT LABORATORIES	69	APC	ANADARKO PETROLEUM CORP
30	DOW	DOW CHEMICAL CO	70	C	CITIGROUP INC
31	LMT	LOCKHEED MARTIN CORP	71	CSCO	CISCO SYSTEMS INC
32	EXC	EXELON CORP	72	BIIB	BIOGEN IDEC INC
33	PFE	PFIZER INC	73	QCOM	QUALCOMM INC
34	EMR	EMERSON ELECTRIC CO	74	GILD	GILEAD SCIENCES INC
35	JNJ	JOHNSON & JOHNSON	75	TWX	TIME WARNER INC NEW
36	MMM	3M CO	76	SBUX	STARBUCKS CORP
37	MRK	MERCK & CO INC NEW	77	ALL	ALLSTATE CORP
38	HAL	HALLIBURTON COMPANY	78	SPG	SIMON PROPERTY GROUP INC NEW
39	RTN	RAYTHEON CO	79	COST	COSTCO WHOLESALE CORP NEW
40	F	FORD MOTOR CO DEL	80	UNH	UNITEDHEALTH GROUP INC

Table 2.1: The list of the stocks studied in the empirical section.

r_t	Mean	Std. dev.	Skew.	Kurt.	KPSS
Top 20					
AAPL	0.000893	0.0305	-2.38	67.1	0.161*
MSFT	0.000615	0.0205	-0.0450	6.58	0.432*
BAC	0.000189	0.0287	-0.344	26.3	0.166*
INTC	0.000485	0.0249	-0.382	6.67	0.222*
CSCO	0.000537	0.0272	-0.111	6.75	0.361*
ORCL	0.000665	0.0290	-0.171	10.7	0.139*
GE	0.000311	0.0188	0.00840	7.89	0.284*
F	0.000251	0.0267	0.366	17.5	0.0810*
PFE	0.000434	0.0174	-0.187	3.87	0.462*
QCOM	0.000766	0.0320	0.471	6.63	0.151*
EMC	0.000527	0.0303	-0.307	7.77	0.213*
WFC	0.000524	0.0239	0.788	25.	0.0846*
JPM	0.000408	0.0251	0.244	11.8	0.0815*
GILD	0.00105	0.0325	0.0262	5.51	0.0281*
XOM	0.000436	0.0154	0.0407	9.12	0.125*
T	0.000259	0.0173	0.0668	5.17	0.0783*
C	5.28e-05	0.0310	-0.442	37.5	0.436*
TXN	0.000539	0.0278	0.147	3.50	0.168*
KO	0.000336	0.0145	0.0136	6.36	0.130*
HPQ	0.000394	0.0250	-0.0381	8.34	0.139*
Summary					
Top 20 - Avg.	0.000484	0.0249	-0.112	14.0	-
Top 20 - Freq.	-	-	-	-	100
All - Avg.	0.000485	0.0215	-0.184	19.3	-
All - Freq.	-	-	-	-	99

Table 2.2: Sample statistics of return time series. "Top 20" corresponds to the twenty stocks of the S&P 100 Index (as of 2014-12-31) with highest mean daily volume (number of traded shares) from 1994-01-03 to 2014-12-31. "All" corresponds to all S&P 100 stocks as of 2014-12-31 (provided that the price history goes back to 1994-01-03). "Avg." denotes mean statistic over a given sample of stocks. "Freq." is the percentage of stocks for wich the KPSS statistical test does not reject stationarity hypothesis at 5% level (* in last column).

$\log V_t/V_{t-1}$					Autoregressive					Moving average						
	Mean	Std. dev.	KPSS	Intercept	1	2	3	4	5	1	2	3	4	5	σ_u	R^2
Top 20																
AAPL	-1.77e-05	0.428	0.00897*	-	0.509	-	-	-	-	-0.938	-	-	-	-	0.383	0.2
MSFT	-0.000101	0.367	0.00851*	-	1.17	-0.260	-	-	-	-1.69	0.699	-	-	-	0.319	0.24
BAC	0.000530	0.362	0.0110*	-	-1.19	-0.0647	0.503	0.0859	-	0.701	-0.682	-0.792	-	-	0.320	0.22
INTC	-0.000154	0.359	0.00666*	-	0.765	-	-	-	-	-1.28	0.202	0.0897	-	-	0.314	0.24
CSCO	-0.000218	0.357	0.00398*	-	0.497	-	-	-	-	-0.953	-	-	-	-	0.316	0.22
ORCL	-0.000346	0.386	0.00290*	-0.000223	1.31	-0.365	-	-	-	-1.81	0.814	-	-	-	0.338	0.23
GE	0.000114	0.333	0.00575*	-	0.211	0.396	-	-	-	-0.752	-0.412	0.196	-	-	0.290	0.24
F	0.000209	0.419	0.00847*	-	1.23	-0.307	-	-	-	-1.74	0.740	-	-	-	0.367	0.23
PFE	0.000122	0.351	0.00932*	-	0.493	0.653	-0.228	-	-	-0.998	-0.606	0.586	0.0218	-	0.308	0.23
QCOM	-1.26e-05	0.447	0.00789*	-	-0.485	-0.109	0.698	-	-	-0.0227	-0.315	-0.920	0.244	0.0649	0.392	0.23
EMC	-3.23e-06	0.417	0.00392*	-	0.603	-	-	-	-	-1.12	0.147	-	-	-	0.365	0.23
WFC	0.000277	0.380	0.0135*	-	0.0655	0.531	0.666	-0.489	0.0487	-0.630	-0.637	-0.434	0.906	-0.188	0.326	0.26
JPM	0.000331	0.369	0.00638*	-	0.833	-	-	-	-	-1.34	0.213	0.115	0.0241	-	0.324	0.23
GILD	0.000523	0.650	0.0129*	-	0.971	-0.0566	-	-	-	-1.61	0.525	0.0604	0.0294	-	0.543	0.3
XOM	0.000105	0.308	0.00767*	-	0.612	-	-	-	-	-1.20	0.235	-	-	-	0.263	0.27
T	0.000534	0.356	0.00567*	0.000483	0.619	-	-	-	-	-1.19	0.213	-	-	-	0.306	0.26
C	0.000549	0.370	0.0133*	-	0.834	-	-	-	-	-1.33	0.265	0.0382	0.0484	-	0.327	0.22
TXN	-6.32e-05	0.379	0.0107*	-	0.654	-	-	-	-	-1.18	0.202	-	-	-	0.331	0.24
KO	0.000104	0.356	0.00514*	-	0.0618	0.469	-	-	-	-0.617	-0.586	0.222	-	-	0.308	0.25
HPQ	6.68e-05	0.402	0.00582*	-	0.197	0.505	-0.0444	-	-	-0.700	-0.625	0.265	0.0746	-	0.351	0.24
Summary																
Top 20 - Avg.	0.000127	0.390	-	0.000130	0.498	0.127	0.319	-0.202	0.0487	-1.02	0.0218	-0.0522	0.193	-0.0616	0.340	0.24
Top 20 - Freq.	-	-	100	10	100	55	25	10	5	100	90	55	35	10	100	-
All - Avg.	0.000113	0.419	-	0.000262	0.563	0.191	0.153	-0.0919	-0.0125	-1.12	0.0312	0.0734	0.131	0.0350	0.360	0.26
All - Freq.	-	-	100	5	100	60	28	10	5	100	89	69	41	15	100	-

Table 2.3: Sample statistics of time series of daily difference in log volume as well as fitted ARMA models. "Top 20" corresponds to the twenty stocks of the S&P 100 Index (as of 2014-12-31) with highest mean daily volume (number of traded shares) from 1994-01-03 to 2014-12-31. "All" corresponds to all S&P 100 stocks as of 2014-12-31 (provided that the price history goes back to 1994-01-03). "Avg." denotes mean parameter over a given sample of stocks. "Freq." is the percentage of stocks for wich: the KPSS statistical test does not reject stationarity hypothesis at 5% level (* in corresponding column), a given parameter of the ARMA specification is selected by the procedure.

1994-01 to 2014-12	Unrestricted	$\gamma_2 = 0$	$\gamma_1 = 0$	$\gamma_1 = \gamma_2$	$\gamma_1 = \gamma_2 = \rho = 0$	$\alpha_1 = \alpha_2 = \beta_1 = 0$	VA-GARCH
AIC							
Unrestricted	-	100	62	99	100	100	100
$\gamma_2 = 0$	0	-	0	0	69	70	79
$\gamma_1 = 0$	38	100	-	99	100	100	100
$\gamma_1 = \gamma_2$	1	100	1	-	100	100	100
$\gamma_1 = \gamma_2 = \rho = 0$	0	31	0	0	-	62	76
$\alpha_1 = \alpha_2 = \beta_1 = 0$	0	30	0	0	38	-	38
VA-GARCH	0	21	0	0	24	62	-
BIC							
Unrestricted	-	100	59	99	100	100	100
$\gamma_2 = 0$	0	-	0	0	59	66	79
$\gamma_1 = 0$	41	100	-	99	100	100	100
$\gamma_1 = \gamma_2$	1	100	1	-	100	100	100
$\gamma_1 = \gamma_2 = \rho = 0$	0	41	0	0	-	62	76
$\alpha_1 = \alpha_2 = \beta_1 = 0$	0	34	0	0	38	-	38
VA-GARCH	0	21	0	0	24	62	-
LRT							
Unrestricted	-	100	62	99	100	100	-
$\gamma_2 = 0$	0	-	-	-	65	-	-
$\gamma_1 = 0$	38	-	-	-	100	-	-
$\gamma_1 = \gamma_2$	1	-	-	-	-	-	-
$\gamma_1 = \gamma_2 = \rho = 0$	0	35	0	-	-	-	-
$\alpha_1 = \alpha_2 = \beta_1 = 0$	0	-	-	-	-	-	-
VA-GARCH	-	-	-	-	-	-	-

Table 2.4: Summary of the model selection procedure according to AIC, BIC and Likelihood Ratio Tests (LRT). In a given panel corresponding to a specific selection criterion, each number represents the percentage of stocks for which the model specified as row name is selected with respect to the model specified as column name. VA-GARCH is the model of Lamoureux and Lastrapes (1990a).

1994-01 to 2014-12	μ	σ	α_1	β_1	α_2	γ_1	γ_2	ρ
Top 20								
AAPL	3.35e-05	0.0313***	0.0406***	0.919***	0.0729***	-1.42***	1.83***	0.0735***
	(0.000229)	(0.00692)	(0.00846)	(0.00552)	(0.00668)	(0.103)	(0.0467)	(0.0145)
MSFT	0.000444***	0.0507***	0.159***	0.839***	2.73e-07	-1.41***	1.91***	0.0424***
	(0.000157)	(0.00351)	(0.00164)	(0.00164)	(0.000119)	(0.112)	(0.0517)	(0.0144)
BAC	0.000607***	0.0206***	0.121***	0.807***	0.117***	-0.716***	1.93***	-0.00396
	(0.000159)	(0.00321)	(0.0172)	(0.0200)	(0.0213)	(0.116)	(0.0518)	(0.0144)
INTC	0.000647***	0.0633***	0.0845***	0.891***	0.0478***	-1.28***	2.07***	0.00360
	(0.000187)	(0.00719)	(0.00811)	(0.00868)	(0.000279)	(0.118)	(0.0543)	(0.0145)
CSCO	0.000552***	0.0707***	0.0133***	0.940***	0.0926***	-1.54***	1.87***	-0.0198
	(0.000161)	(0.00621)	(0.00272)	(0.00272)	(2.00e-08)	(0.119)	(0.0532)	(0.0140)
ORCL	0.000331*	0.121***	0.124***	0.862***	0.0277***	-1.83***	1.63***	0.00488
	(0.000200)	(0.00880)	(0.00737)	(0.00737)	(2.00e-08)	(0.113)	(0.0482)	(0.0143)
GE	4.38e-05	0.0336***	0.122***	0.875***	1.44e-07***	-0.431***	0.610***	-0.0793***
	(0.000164)	(0.00610)	(0.0208)	(0.0207)	(2.00e-08)	(0.142)	(0.0514)	(0.0137)
F	0.000449*	0.0292	0.0621**	0.919***	0.0243	-1.03***	0.968***	-0.0141
	(0.000258)	(0.0180)	(0.0248)	(0.0192)	(0.0151)	(0.109)	(0.0260)	(0.0143)
PFE	0.000314*	0.0183***	0.134***	0.834***	0.0325	-0.942***	1.80***	0.0105
	(0.000161)	(0.00210)	(0.0283)	(0.0269)	(0.0213)	(0.121)	(0.0558)	(0.0143)
QCOM	0.000453**	0.0546***	0.00439	0.947***	0.0962***	-2.14***	1.25***	-0.00369
	(0.000214)	(0.00653)	(0.00306)	(0.00306)	(2.00e-08)	(0.0930)	(0.0414)	(0.0138)
EMC	0.000343	0.148***	0.146***	0.854***	3.38e-07***	-0.523***	1.86***	0.0213
	(0.000212)	(0.0110)	(0.00798)	(0.00798)	(2.00e-08)	(0.0998)	(0.0486)	(0.0145)
WFC	0.000520***	0.0308***	0.134***	0.824***	0.0771***	-0.784***	1.28***	0.00429
	(0.000142)	(0.00256)	(0.0172)	(0.0181)	(0.0129)	(0.104)	(0.0435)	(0.0139)
JPM	-6.30e-05	0.0211***	0.0708***	0.773***	0.259***	-0.232*	1.88***	0.0226
	(0.000187)	(0.00191)	(0.0181)	(0.0464)	(0.0504)	(0.121)	(0.0542)	(0.0140)
GILD	0.000155	0.0336***	0.00419	0.956***	0.0746***	-0.353***	1.09***	0.0707***
	(0.000237)	(0.00244)	(0.00320)	(0.00266)	(0.00378)	(0.0523)	(0.0298)	(0.0144)
XOM	0.000324**	0.0135***	0.0294***	0.928***	0.0658***	-1.07***	1.98***	-0.00371
	(0.000144)	(0.00119)	(0.00717)	(0.00712)	(0.0129)	(0.122)	(0.0648)	(0.0142)
T	0.000370***	0.0600***	0.0955***	0.875***	0.0578***	-0.965***	1.85***	-0.00963
	(0.000134)	(0.00616)	(0.00498)	(0.00551)	(0.000171)	(0.111)	(0.0559)	(0.0144)
C	-2.20e-06	0.0259***	0.0883***	0.781***	0.234***	-0.889***	1.75***	0.0156
	(0.000179)	(0.00367)	(0.0147)	(0.0313)	(0.0369)	(0.114)	(0.0487)	(0.0140)
TXN	0.000783***	0.0442***	0.118***	0.880***	2.56e-07***	-1.10***	1.78***	0.0257*
	(0.000208)	(0.00734)	(0.0193)	(0.0193)	(2.00e-08)	(0.115)	(0.0521)	(0.0144)
KO	0.000277**	0.0264***	0.0431***	0.935***	0.0417***	-0.998***	1.97***	0.0189
	(0.000110)	(0.00289)	(0.00551)	(0.00600)	(0.000157)	(0.113)	(0.0547)	(0.0144)
HPQ	0.000102	0.0749***	0.159***	0.821***	0.0363***	-1.06***	1.62***	0.0239
	(0.000201)	(0.0151)	(0.0286)	(0.0286)	(0.000804)	(0.110)	(0.0304)	(0.0147)
Summary								
Top 20 - Avg.	0.000334	0.0486	0.0877	0.873	0.0679	-1.04	1.65	0.0102
Top 20 - Freq.	65	95	90	100	85	100	100	25
All - Avg.	0.000263	0.0377	0.0815	0.870	0.0734	-0.852	1.56	0.0234
All - Freq.	56	96	95	100	86	100	100	52

Table 2.5: VF-GARCH model fitted parameters from 1994-01-03 to 2014-12-31. The numbers in parentheses are standard errors. "Top 20" corresponds to the twenty stocks of the S&P 100 Index (as of 2014-12-31) with highest mean daily volume (number of traded shares) from 1994-01-03 to 2014-12-31. "All" corresponds to all S&P 100 stocks as of 2014-12-31 (provided that the price history goes back to 1994-01-03). "Avg." denotes mean parameter over a given sample of stocks. "Freq." is the percentage of stocks for wich a given parameter is significant.

1994-01 to 2014-12	μ	σ	α_1	β_1	α_2
Top 20					
AAPL	0.00167***	0.0530***	0.0541***	0.906***	0.0719***
	(0.000310)	(0.00538)	(0.00684)	(0.00650)	(0.00777)
MSFT	0.000607***	0.0204***	0.0308***	0.937***	0.0434***
	(0.000221)	(0.00134)	(0.00542)	(0.00773)	(0.00967)
BAC	0.000416**	0.0180***	0.0202***	0.933***	0.0744***
	(0.000204)	(0.00121)	(0.00531)	(0.00937)	(0.0106)
INTC	0.000669**	0.0239***	0.0247***	0.897***	0.110***
	(0.000265)	(0.00126)	(0.00580)	(0.0163)	(0.0229)
CSCO	0.000683**	0.0251***	0.00597*	0.958***	0.0564***
	(0.000285)	(0.00150)	(0.00361)	(0.00439)	(0.00818)
ORCL	0.000686***	0.0364***	0.0314***	0.938***	0.0551***
	(0.000266)	(0.00427)	(0.00565)	(0.00481)	(0.00896)
GE	0.000486***	0.0167***	0.0229***	0.943***	0.0559***
	(0.000179)	(0.00156)	(0.00534)	(0.00739)	(0.00954)
F	4.59e-05	0.0259***	0.0494***	0.939***	0.00152***
	(0.000282)	(0.00164)	(0.00832)	(0.0100)	(2.00e-08)
PFE	0.000345*	0.0217***	0.0425***	0.938***	0.0299***
	(0.000195)	(0.00328)	(0.00740)	(0.00670)	(0.0112)
QCOM	0.000778***	0.0433***	0.0430***	0.934***	0.0412***
	(0.000288)	(0.00319)	(0.00603)	(0.00578)	(0.00401)
EMC	0.000778***	0.0334***	0.0116***	0.963***	0.0475***
	(0.000272)	(0.00332)	(0.00322)	(0.00322)	(2.00e-08)
WFC	0.000440***	0.0238***	0.0170***	0.933***	0.0959***
	(0.000160)	(0.00270)	(0.00451)	(0.00407)	(0.00674)
JPM	0.000542**	0.0235***	0.0190***	0.943***	0.0692***
	(0.000210)	(0.00375)	(0.00483)	(0.00547)	(0.0109)
GILD	0.000965***	0.0341***	0.0170***	0.975***	0.0118***
	(0.000331)	(0.00403)	(0.00257)	(0.00257)	(0.00100)
XOM	0.000429**	0.0141***	0.0352***	0.917***	0.0616***
	(0.000168)	(0.000729)	(0.00673)	(0.00855)	(0.0113)
T	0.000328*	0.0195***	0.0447***	0.934***	0.0350***
	(0.000175)	(0.00335)	(0.00727)	(0.00733)	(0.0108)
C	0.000430**	0.0304***	0.0393***	0.932***	0.0524***
	(0.000212)	(0.00312)	(0.00592)	(0.00561)	(0.00482)
TXN	0.000706***	0.0307**	0.0174***	0.955***	0.0516**
	(0.000200)	(0.0140)	(0.00294)	(0.00642)	(0.0241)
KO	0.000488***	0.0156***	0.0272***	0.934***	0.0652***
	(0.000149)	(0.00188)	(0.00605)	(0.00673)	(0.00973)
HPQ	0.000761***	0.0256***	0.00796	0.936***	0.0808**
	(0.000287)	(0.00168)	(0.00542)	(0.0322)	(0.0370)
Summary					
Top 20 - Avg.	0.000613	0.0268	0.0281	0.937	0.0555
Top 20 - Freq.	95	100	95	100	100
All - Avg.	0.000569	0.0247	0.0350	0.923	0.0623
All - Freq.	90	100	96	100	99

Table 2.6: GJR-GARCH model fitted parameters from 1994-01-03 to 2014-12-31. The numbers in parentheses are standard errors. "Top 20" corresponds to the twenty stocks of the S&P 100 Index (as of 2014-12-31) with highest mean daily volume (number of traded shares) from 1994-01-03 to 2014-12-31. "All" corresponds to all S&P 100 stocks as of 2014-12-31 (provided that the price history goes back to 1994-01-03). "Avg." denotes mean parameter over a given sample of stocks. "Freq." is the percentage of stocks for wich a given parameter is significant.

1994-01 to 2014-12	GARCH contrib.	Volume contrib.	Exp. volume contrib.	Unexp. volume contrib.
Top 20				
AAPL	0.49	0.51	0.066	0.45
MSFT	0.59	0.41	0.058	0.35
BAC	0.75	0.25	0.0094	0.24
INTC	0.60	0.40	0.041	0.36
CSCO	0.67	0.33	0.049	0.28
ORCL	0.68	0.32	0.087	0.23
GE	0.95	0.047	0.0061	0.041
F	0.74	0.26	0.064	0.19
PFE	0.59	0.41	0.031	0.38
QCOM	0.70	0.30	0.14	0.16
EMC	0.65	0.35	0.0081	0.34
WFC	0.85	0.15	0.017	0.13
JPM	0.70	0.30	0.0013	0.30
GILD	0.70	0.30	0.013	0.29
XOM	0.56	0.44	0.042	0.39
T	0.69	0.31	0.027	0.28
C	0.75	0.25	0.016	0.23
TXN	0.67	0.33	0.034	0.30
KO	0.62	0.38	0.029	0.35
HPQ	0.61	0.39	0.044	0.34
Summary				
Top 20 - Avg.	0.68	0.32	0.039	0.28
Top 20 - Freq. $> 50\%$	95	5	0	0
All - Avg.	0.64	0.36	0.035	0.32
All - Freq. $> 50\%$	92	8	0	4

Table 2.7: Components' contribution to variations in stock return variance. In the first panel, for each row, the numbers correspond respectively to the ratios $\mathrm{Var}\left(\log \tau_t\right)/\mathrm{Var}\left(\log \tau_t g_t\right)$, $\mathrm{Var}\left(\log g_t\right)/\mathrm{Var}\left(\log \tau_t g_t\right)$, $\mathrm{Var}\left(\gamma_1 v_t^e\right)/\mathrm{Var}\left(\log \tau_t g_t\right)$, and $\mathrm{Var}\left(\gamma_2 v_t^u\right)/\mathrm{Var}\left(\log \tau_t g_t\right)$. "Top 20" corresponds to the twenty stocks of the S&P 100 Index (as of 2014-12-31) with highest mean daily volume (number of traded shares) from 1994-01-03 to 2014-12-31. "All" corresponds to all S&P 100 stocks as of 2014-12-31 (provided that the price history goes back to 1994-01-03). "Avg." denotes mean ratio over a given sample of stocks. "Freq. $> 50\%$" is the percentage of stocks for which a given component represent more than half of the variability of the variance estimates.

1994-01 to 2001-02	μ	σ	α_1	β_1	α_2	γ_1	γ_2	ρ
Top 20 - Avg.	-0.000191	0.0283	0.0809	0.829	0.0895	-0.750	1.46	0.0688
Top 20 - Freq.	35	100	100	100	85	90	100	75
All - Avg.	-0.000181	0.0236	0.0876	0.835	0.0631	-0.592	1.32	0.0811
All - Freq.	22	99	92	100	68	91	100	88
2001-03 to 2001-11	μ	σ	α_1	β_1	α_2	γ_1	γ_2	ρ
Top 20 - Avg.	-0.000860	0.0421	0.0665	0.838	0.124	-0.618	1.71	-0.0659
Top 20 - Freq.	10	100	25	100	55	25	100	40
All - Avg.	-0.000199	0.0285	0.0785	0.810	0.0922	-0.619	1.62	-0.0568
All - Freq.	9	100	34	95	51	29	100	34
2001-12 to 2007-11	μ	σ	α_1	β_1	α_2	γ_1	γ_2	ρ
Top 20 - Avg.	1.82e-05	0.0273	0.0724	0.879	0.0724	-0.989	1.92	0.0177
Top 20 - Freq.	25	100	50	100	70	95	100	55
All - Avg.	0.000169	0.0222	0.0654	0.874	0.0785	-0.865	1.83	0.0150
All - Freq.	35	98	70	100	75	95	100	36
2007-12 to 2009-06	μ	σ	α_1	β_1	α_2	γ_1	γ_2	ρ
Top 20 - Avg.	-0.000844	0.0362	0.105	0.817	0.103	-1.88	1.99	-0.0323
Top 20 - Freq.	15	100	50	100	85	90	100	25
All - Avg.	-0.000470	0.0284	0.0797	0.828	0.107	-1.32	1.81	-0.0444
All - Freq.	10	99	45	99	68	80	100	31
2009-07 to 2014-12	μ	σ	α_1	β_1	α_2	γ_1	γ_2	ρ
Top 20 - Avg.	0.000708	0.0195	0.0646	0.834	0.125	-1.25	2.29	-0.0325
Top 20 - Freq.	75	100	65	100	85	100	100	30
All - Avg.	0.000693	0.0218	0.0743	0.839	0.102	-1.10	2.04	-0.0248
All - Freq.	74	98	58	100	72	99	100	35

Table 2.8: Summary of VF-GARCH model fitted parameters over five sub-samples. The five panels correspond respectively to the sub-periods: 1) January 1994 to February 2001, 2) Mars 2001 to November 2001, 3) December 2001 to November 2007, 4) December 2007 to June 2009, and 5) July 2009 to December 2014. "Top 20" corresponds to the twenty stocks of the S&P 100 Index (as of 2014-12-31) with highest mean daily volume (number of traded shares) from 1994-01-03 to 2014-12-31. "All" corresponds to all S&P 100 stocks as of 2014-12-31 (provided that the price history goes back to 1994-01-03). "Avg." denotes mean parameter over a given sample of stocks. "Freq." is the percentage of stocks for wich a given parameter is significant.

1994-01 to 2001-02	μ	σ	α_1	β_1	α_2
Top 20 - Avg.	0.00111	0.0284	0.0402	0.881	0.0687
Top 20 - Freq.	75	100	80	100	85
All - Avg.	0.000914	0.0239	0.0413	0.892	0.0551
All - Freq.	66	100	81	100	75
2001-03 to 2001-11	μ	σ	α_1	β_1	α_2
Top 20 - Avg.	-0.00140	0.0448	0.0477	0.827	0.137
Top 20 - Freq.	5	95	40	100	70
All - Avg.	-0.000664	0.0354	0.0469	0.787	0.156
All - Freq.	1	96	44	95	58
2001-12 to 2007-11	μ	σ	α_1	β_1	α_2
Top 20 - Avg.	0.000333	0.0246	0.0278	0.934	0.0466
Top 20 - Freq.	25	100	85	100	85
All - Avg.	0.000394	0.0220	0.0370	0.895	0.0669
All - Freq.	38	99	74	99	80
2007-12 to 2009-06	μ	σ	α_1	β_1	α_2
Top 20 - Avg.	-0.000942	0.0325	0.0321	0.872	0.137
Top 20 - Freq.	10	100	65	100	90
All - Avg.	-0.000692	0.0291	0.0329	0.867	0.120
All - Freq.	5	100	42	99	88
2009-07 to 2014-12	μ	σ	α_1	β_1	α_2
Top 20 - Avg.	0.000556	0.0217	0.0356	0.873	0.0808
Top 20 - Freq.	40	100	75	100	85
All - Avg.	0.000589	0.0241	0.0560	0.850	0.0958
All - Freq.	59	98	65	99	84

Table 2.9: Summary of GJR-GARCH model ($\gamma_1 = \gamma_2 = \rho = 0$) fitted parameters over five sub-samples. The five panels correspond respectively to the sub-periods: 1) January 1994 to February 2001, 2) Mars 2001 to November 2001, 3) December 2001 to November 2007, 4) December 2007 to June 2009, and 5) July 2009 to December 2014. "Top 20" corresponds to the twenty stocks of the S&P 100 Index (as of 2014-12-31) with highest mean daily volume (number of traded shares) from 1994-01-03 to 2014-12-31. "All" corresponds to all S&P 100 stocks as of 2014-12-31 (provided that the price history goes back to 1994-01-03). "Avg." denotes mean parameter over a given sample of stocks. "Freq." is the percentage of stocks for wich a given parameter is significant.

1994-01 to 2001-02	GARCH contrib.	Volume contrib.	Exp. volume contrib.	Unexp. volume contrib.
Top 20 - Avg.	0.42	0.58	0.053	0.52
Top 20 - Freq. $> 50\%$	25	75	0	65
All - Avg.	0.44	0.56	0.043	0.51
All - Freq. $> 50\%$	35	65	0	58
2001-03 to 2001-11	GARCH contrib.	Volume contrib.	Exp. volume contrib.	Unexp. volume contrib.
Top 20 - Avg.	0.42	0.58	0.036	0.54
Top 20 - Freq. $> 50\%$	35	65	5	65
All - Avg.	0.31	0.69	0.062	0.63
All - Freq. $> 50\%$	19	81	0	72
2001-12 to 2007-11	GARCH contrib.	Volume contrib.	Exp. volume contrib.	Unexp. volume contrib.
Top 20 - Avg.	0.57	0.43	0.037	0.39
Top 20 - Freq. $> 50\%$	70	30	0	30
All - Avg.	0.49	0.51	0.038	0.47
All - Freq. $> 50\%$	49	51	0	44
2007-12 to 2009-06	GARCH contrib.	Volume contrib.	Exp. volume contrib.	Unexp. volume contrib.
Top 20 - Avg.	0.54	0.46	0.078	0.39
Top 20 - Freq. $> 50\%$	45	55	0	25
All - Avg.	0.58	0.42	0.060	0.36
All - Freq. $> 50\%$	59	41	0	18
2009-07 to 2014-12	GARCH contrib.	Volume contrib.	Exp. volume contrib.	Unexp. volume contrib.
Top 20 - Avg.	0.33	0.67	0.059	0.61
Top 20 - Freq. $> 50\%$	5	95	0	85
All - Avg.	0.37	0.63	0.057	0.57
All - Freq. $> 50\%$	25	75	0	66

Table 2.10: Summary of volume contribution to variations in stock return variance, over five sub-samples. The five panels correspond respectively to the sub-periods: 1) January 1994 to February 2001, 2) Mars 2001 to November 2001, 3) December 2001 to November 2007, 4) December 2007 to June 2009, and 5) July 2009 to December 2014. "Top 20" corresponds to the twenty stocks of the S&P 100 Index (as of 2014-12-31) with highest mean daily volume (number of traded shares) from 1994-01-03 to 2014-12-31. "All" corresponds to all S&P 100 stocks as of 2014-12-31 (provided that the price history goes back to 1994-01-03). "Avg." denotes mean ratio over a given sample of stocks. "Freq. $> 50\%$" is the percentage of stocks for which a given component represent more than half of the variability of the variance estimate.

IBM	Price change (%)	Change in variance (%)	Change in τ_t (%)	Change in g_t (%)	Unexpected volume
2000-07-18	-2.4	-8.42	-10.54	2.38	0.05
2000-07-19	0.54	11.5	10.8	0.63	0.17
2000-07-20	**12.26**	**256.54**	**-12.79**	**308.85**	**1.25**
2000-07-21	-2.21	-48.83	43.28	-64.29	-0.2
2000-07-24	-1.71	-48.44	-4.05	-46.26	-0.07
	Price change (%)	Change in variance (%)	Change in τ_t (%)	Change in g_t (%)	Unexpected volume
2000-10-16	1.87	-17.9	-5.08	-13.5	0.13
2000-10-17	1.67	-12.07	-14.7	3.08	0.2
2000-10-18	**-16.89**	**272.69**	**-14.52**	**335.97**	**1.36**
2000-10-19	1.04	16.02	87.29	-38.05	0.32
2000-10-20	-1.77	-73.25	-16.53	-67.96	-0.2
	Price change (%)	Change in variance (%)	Change in τ_t (%)	Change in g_t (%)	Unexpected volume
2001-01-16	-1.14	-40.16	-13.8	-30.57	-0.37
2001-01-17	4.16	94.16	-0.74	95.6	0.26
2001-01-18	**11.35**	**337.6**	**16.25**	**276.42**	**1.02**
2001-01-19	2.68	-23.94	36.85	-44.42	0.05
2001-01-22	-2.45	-57.09	-11.39	-51.57	-0.16

Table 2.11: The impact of earning surprises on volatility. The table reports the variation of price and variance as well as unexpected volumes during three periods of earning surprises for IBM. The dates of earning surprises are recorded in bold.

IBM	Price change (%)	Effect on τ_{t+10} (×100)	Effect on τ_{t+20} (×100)	Effect on τ_{t+30} (×100)
2000-07-18	-2.4	0.011	0.0015	0.00020
2000-07-19	0.54	0.00016	2.2e-05	3.1e-06
2000-07-20	**12.26**	**0.024**	**0.0033**	**0.00046**
2000-07-21	-2.21	0.0063	0.00086	0.00012
2000-07-24	-1.71	0.0069	0.00095	0.00013
	Price change (%)	Effect on τ_{t+10} (×100)	Effect on τ_{t+20} (×100)	Effect on τ_{t+30} (×100)
2000-10-16	1.87	0.0019	0.00026	3.6e-05
2000-10-17	1.67	0.0015	0.00020	2.8e-05
2000-10-18	**-16.89**	**0.099**	**0.014**	**0.0019**
2000-10-19	1.04	0.00018	2.5e-05	3.4e-06
2000-10-20	-1.77	0.0060	0.00082	0.00011
	Price change (%)	Effect on τ_{t+10} (×100)	Effect on τ_{t+20} (×100)	Effect on τ_{t+30} (×100)
2001-01-16	-1.14	0.0062	0.00086	0.00012
2001-01-17	4.16	0.015	0.0020	0.00028
2001-01-18	**11.35**	**0.029**	**0.0040**	**0.00056**
2001-01-19	2.68	0.0028	0.00039	5.3e-05
2001-01-22	-2.45	0.014	0.0019	0.00027

Table 2.12: The impact of returns on τ_t component. The table reports the variation of price and impact on τ_t the following days, during three specific periods of earning surprises for IBM. The exact dates of earning surprises are recorded in bold.

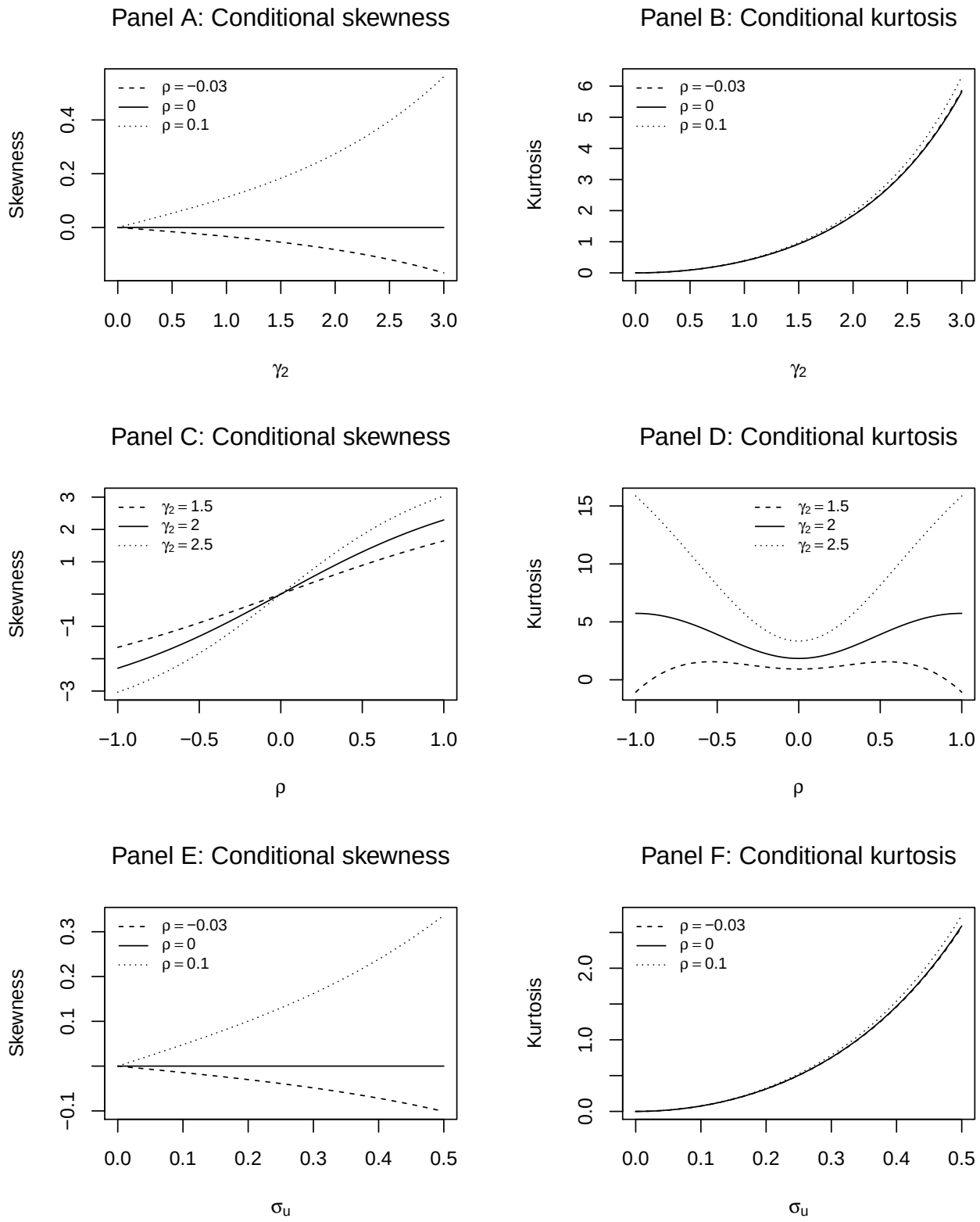

Figure 2.1: Skewness and (excess) kurtosis of returns as functions of γ_2, ρ and σ_u. The Figure presents the analytic values of skewness and kurtosis of returns calculated using the closed form expressions of the normal log-normal moments. Each panel corresponds to a given moment as a function of a specific parameter. All other parameters are constant and set to an arbitrary plausible value according to our empirical results. Except in panels E and F, σ_u is set to 0.35, which corresponds to the median of fitted values over all stocks in our sample from 1994-01-03 to 2014-12-31.

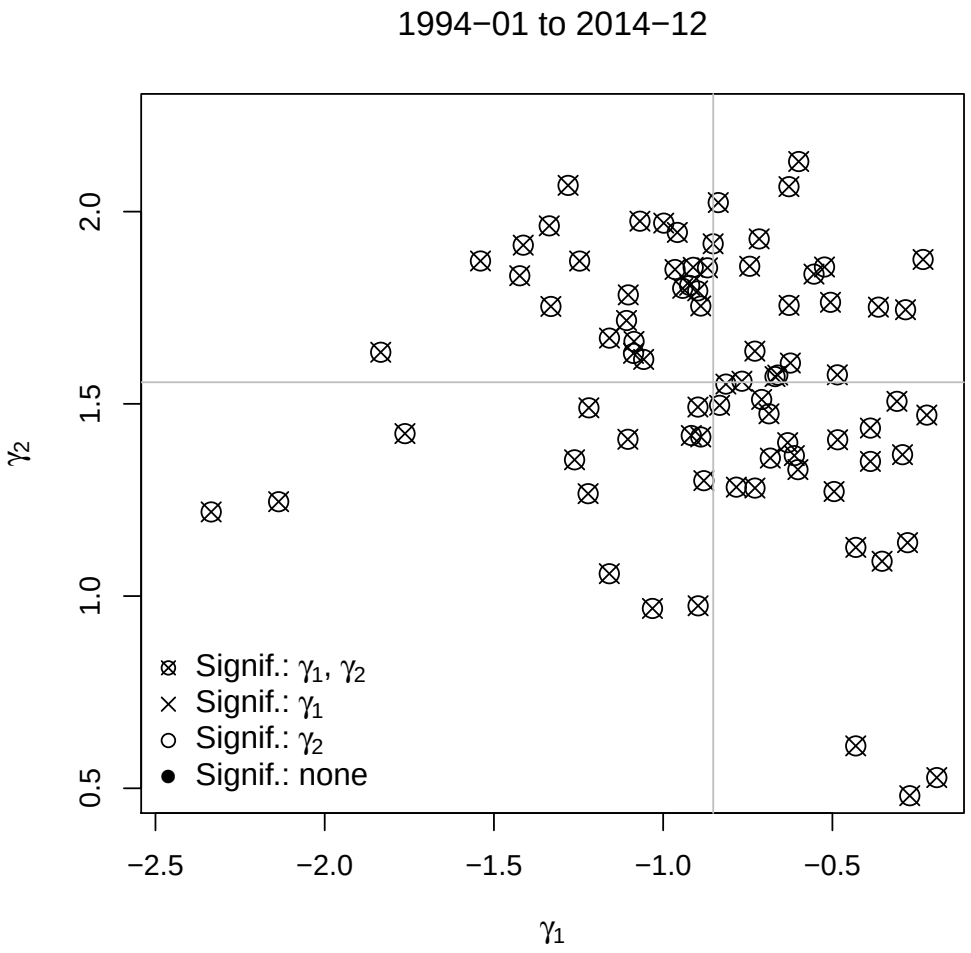

Figure 2.2: Scatter plot of VF-GARCH fitted parameters: γ_1 vs. γ_2, i.e., expected volume impact vs. unexpected volume impact. The Figure provides the scatter-plot of parameter pairs (γ_1, γ_2) for the period 1994-2014. The vertical and horizontal grey lines correspond respectively to the average value over all stocks of γ_1 and γ_2.

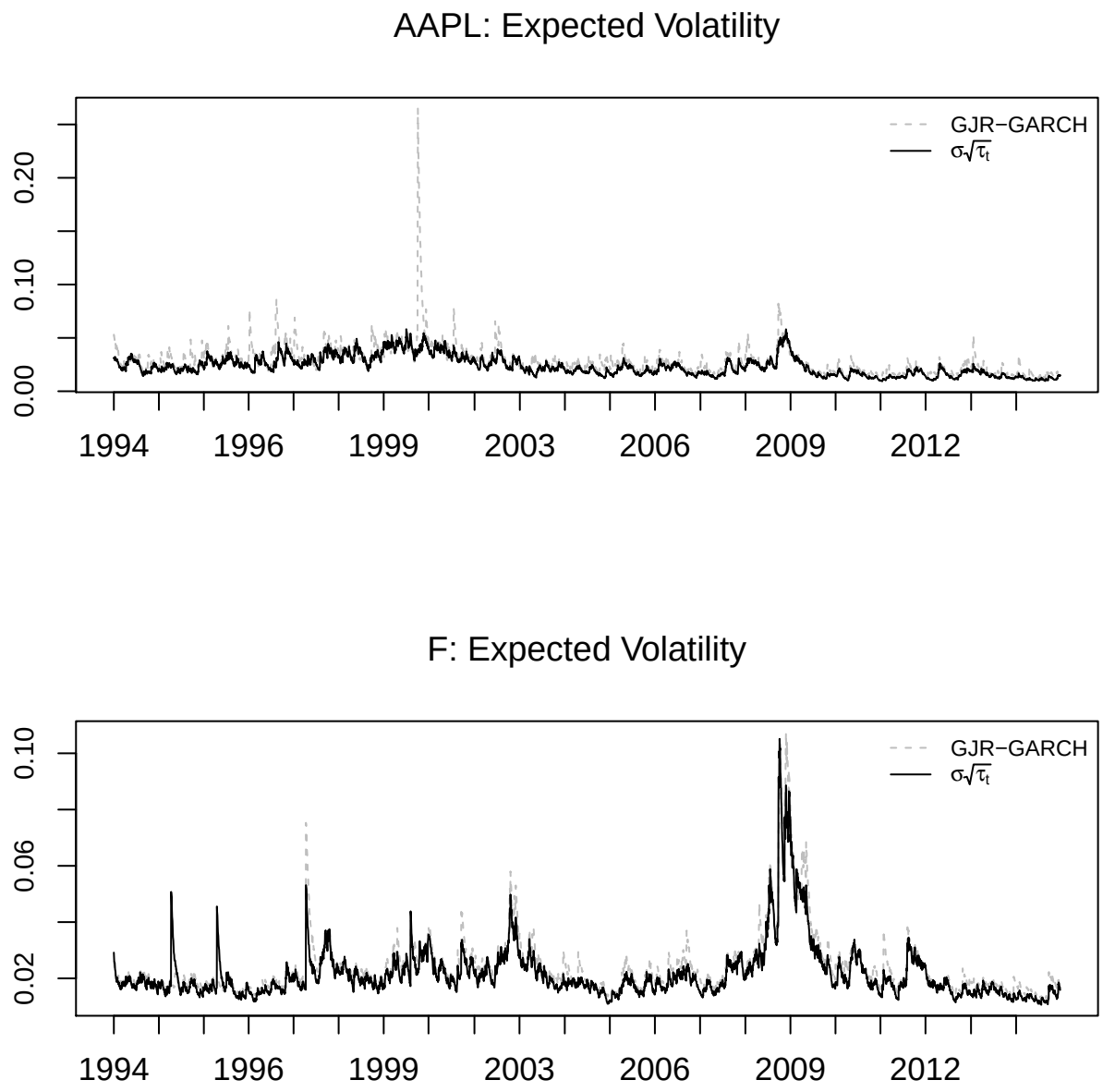

Figure 2.3: AAPL (Apple Inc.) and F (Ford Motor Co Del) filtered volatility component $\sigma\sqrt{\tau_t}$ and GJR-GARCH volatility forecast. The sample begins on 1994-01-03 and ends on 2014-12-31. Both panels show the predicted volatility from the GJR-GARCH model (dashed grey line) and the VF-GARCH model's filtered component of volatility, unrelated to trading volume (continuous black line).

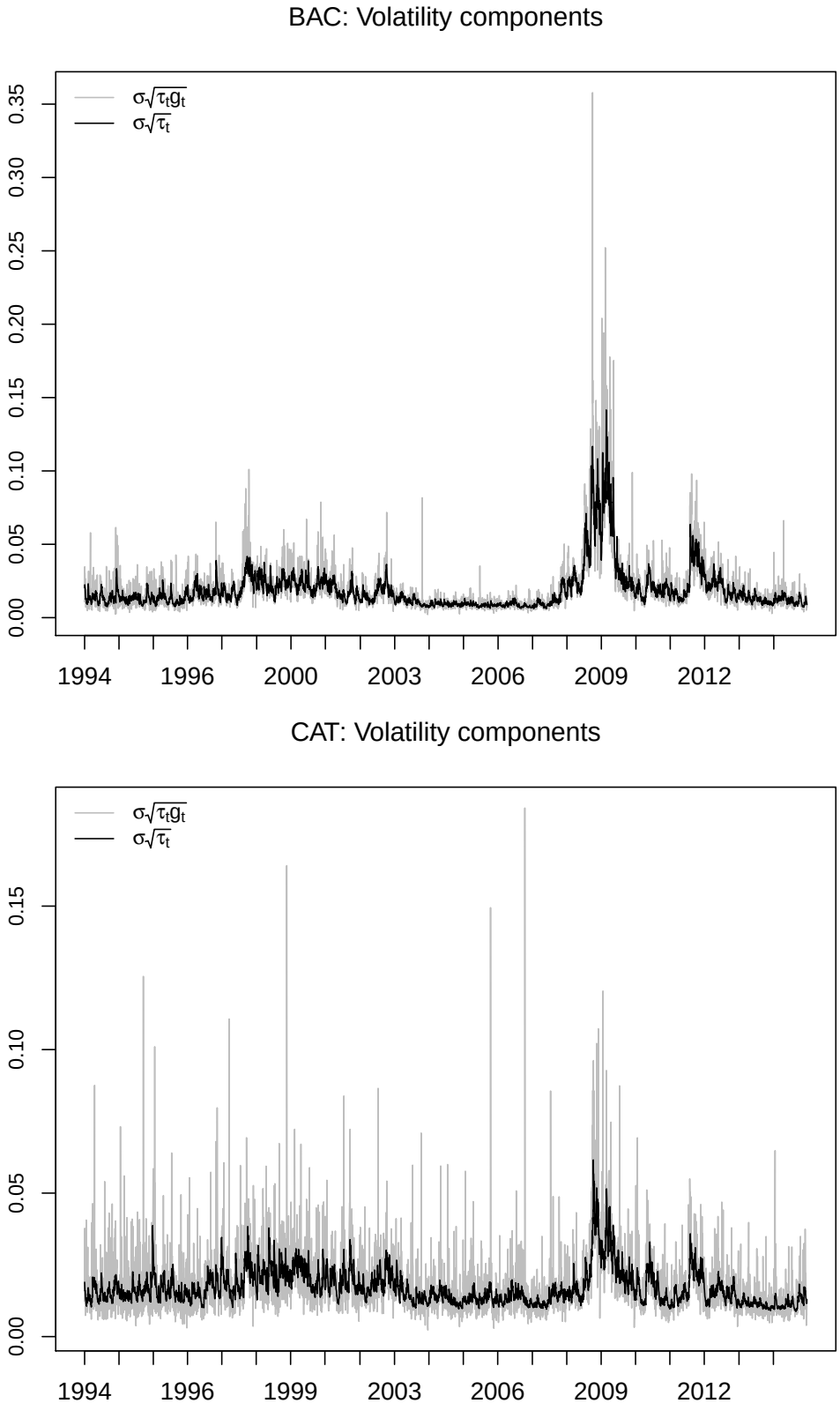

Figure 2.4: BAC (Bank of America Corp.) and CAT (Caterpillar Inc.) volatility estimate components: total volatility (grey) and filtered persistent component unrelated to volume (black). The sample begins on 1994-01-03 and ends on 2014-12-31. Both panels show the volatility components from the VF-GARCH model.

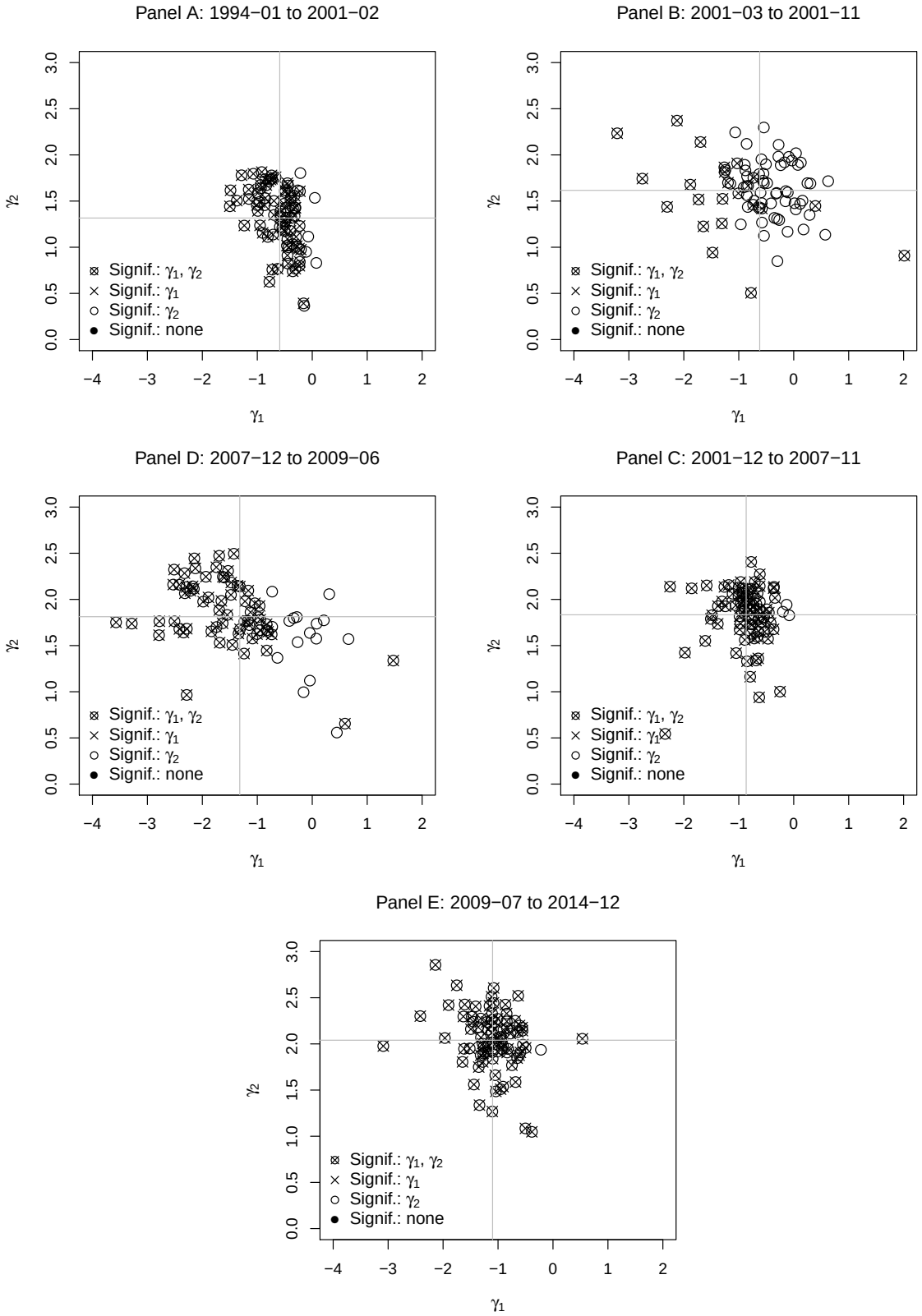

Figure 2.5: Scatter plot of VF-GARCH fitted parameters: γ_1 vs. γ_2, i.e., expected volume impact vs. unexpected volume impact. Each panel corresponds to a given sample. The Figure provides the scatter-plot of parameter pairs (γ_1, γ_2). The vertical and horizontal grey lines correspond respectively to the average value over all stocks of γ_1 and γ_2.

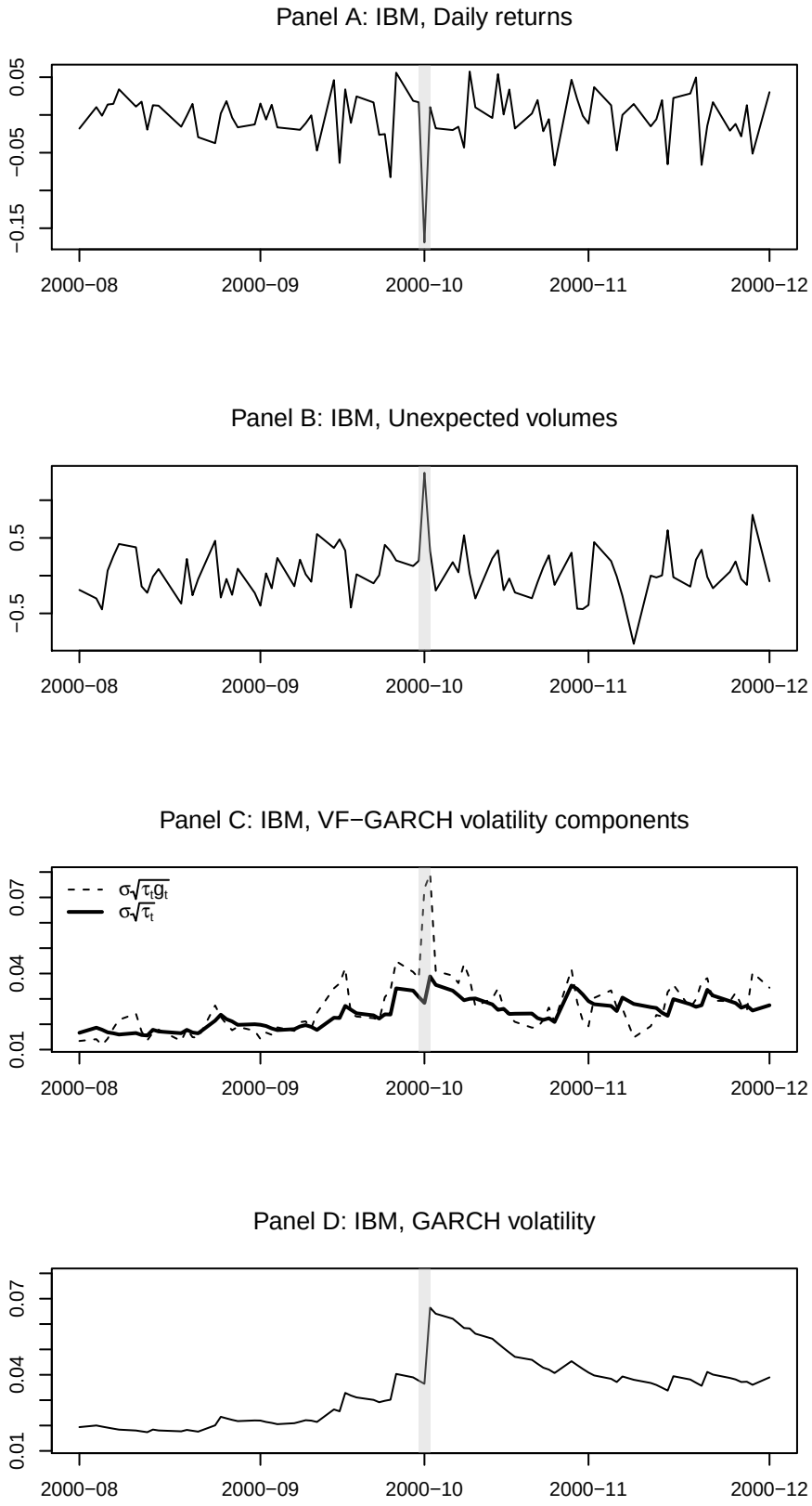

Figure 2.6: The evolution of returns, unexpected volumes, VF-GARCH volatility components and GJR-GARCH estimates during an earning surprise for IBM. Panel A plots the daily returns, Panels B displays the unexpected volumes, Panel C presents the evolution of the total VF-GARCH volatility (dotted line) and filtered GARCH component (bold line), and Panel D plots the GJR-GARCH ($\gamma_1 = \gamma_2 = \rho = 0$). The grayed area corresponds to the earning surprise date.

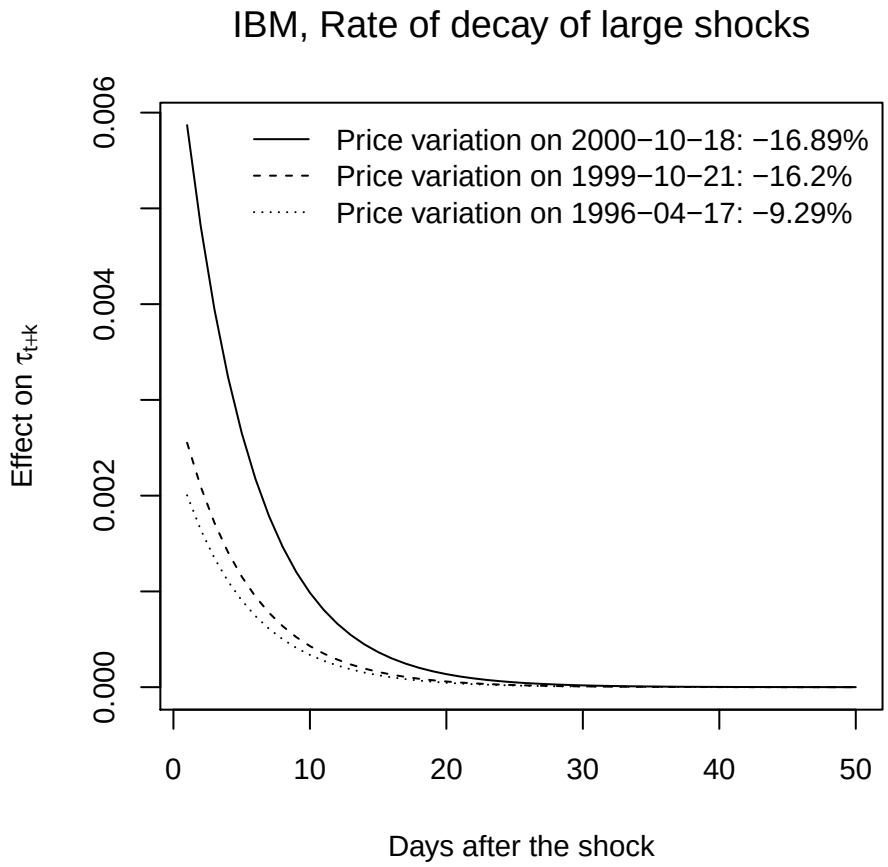

Figure 2.7: The impact of large negative returns on future values of τ_t, for IBM. The different lines correspond to the three largest negative shocks in the price of IBM stock, over the period 1994-01 to 2001-02. Each line presents the impact of the shock on the following values of the component τ_{t+k}: $\beta_1^{k-1}(\alpha_1 + \alpha_2 \mathbb{1}_{r_t<\mu_t}) \frac{(r_t-\mu_t)^2}{\sigma^2 g_t}$ for $1 \leq k \leq 50$.

Chapter 3

Portfolio Management with Drawdown Constraint: An Analysis of Optimal Investment

We analyze optimal investment strategies under the drawdown constraint that the wealth process never falls below a fixed fraction of its running maximum. We derive optimal allocation programs by solving numerically the Hamilton-Jacobi-Bellman equation that characterizes the finite horizon expected utility maximization problem, for investors with power utility as well as S-shape utility. Using stochastic simulations, we find that implementing the drawdown constraint can be gainful in optimal portfolios for the power utility, for some market configurations and investment horizons. However, our study reveals different results in a prospect theory context. According to utility maximization, the optimal strategy with drawdown constraint is not the preferred investment for the S-shape utility investor, who rather prefers the equivalent optimal strategy without constraint. Indeed, the latter investment being similar to a partial portfolio insurance, the additional drawdown constraint does not appear valuable for this investor in optimal portfolios.

Keywords: portfolio optimization, drawdown constraint, prospect theory, behavioral finance.

3.1 Introduction

Portfolio insurance strategies can appear attractive because they are designed to protect assets under management while allowing some upside potential. The risk limitation objective of such strategies can take different forms in terms of capital protection. The most popular allocation programs are the Option Based Portfolio Insurance (OBPI) of Leland and Rubinstein (1976), insuring a given amount at maturity, and the Constant Proportion Portfolio Insurance (CPPI) of Black and Rouhani (1989); Black and Perold (1992). A second well known constraint considered in portfolio insurance strategies is the maximum cumulated loss limitation, namely the Time Invariant Portfolio Protection (TIPP) or maximum drawdown control, studied by Estep and Kritzman (1988); Grossman and Zhou (1993); Cvitanic and Karatzas (1995); Klass and Nowicki (2005). Such capital protection can usually be achieved through a dynamic allocation between a risk-free (reserve) asset and a risky (performance-seeking) asset.

Given the extensive use of portfolio insurance strategies in the asset management industry, numerous papers focus on their effect on financial markets (see Brennan and Schwartz, 1989; Basak, 1995; Grossman and Zhou, 1996, for instance). On the other hand, several authors examine the theoretical reasons of such popularity among market participants. In particular, Leland (1980) studies the nature of investors who will benefit from purchasing a capital protection at maturity, while Brennan, Solanki, et al. (1981) characterize optimal investment portfolio insurance contracts. Given the difficulty to justify the use of portfolio insurance for investors with standard concave utility functions, several authors consider their attractiveness for loss averse investors in a prospect theory context. Indeed, loss aversion can have several effects in the economy and thus explains puzzling behaviors (see in particular Benartzi and Thaler, 1995; Barberis, Huang, and Santos, 2001; Gomes, 2005; Haigh and List, 2005; Grinblatt and Han, 2005; Grüne and Semmler, 2008). The main reason is that, unlike standard utility, the utility function of prospect theory investors is assumed asymmetric: convex for wealth below the reference point (leading to a risk-seeking behavior) and concave for wealth above, i.e., S-shape.

In a behavioral finance context, portfolio insurance strategies have been studied by Vrecko and Branger (2009); Dierkes, Erner, and Zeisberger (2010); Dichtl and Drobetz (2011); Ebert, Koos, and Schneider (2012); Khuman, Constantinou, and Phelps (2012). These studies indicate that prospect theory investors with S-shape utility could prefer portfolio insurance strategies rather than other standard investment solutions, and could profit more than other investors from this allocation method. However, the latter papers do not examine portfolio insurance strategies with optimal investment regarding utility maximization. More specifically, the historical and Monte Carlo simulations of interest for portfolio insurance strategies are run with constant multipliers, i.e., the key parameter driving the exposure to the risky asset, over the investment period. Hence, the resulting allocation can be far from optimal in terms of utility maximization[1].

[1]In addition, in the asset management industry, the exposure to the risky asset is permanently controlled through a time-varying multiplier (based on extreme gap risk estimation for instance, see Cont and Tankov, 2009; Hamidi, Maillet, and Prigent, 2014).

Optimal portfolio choices under loss aversion have been studied by Benartzi and Thaler (1995), Barberis et al. (2001), Berkelaar, Kouwenberg, and Post (2004), Jin and Yu Zhou (2008), Bernard and Ghossoub (2010), and He and Zhou (2011). Carassus, Rásonyi, and Rodrigues (2015) also consider non-concave utility maximization in discrete time. In particular, Berkelaar et al. (2004) show analytically that, without imposing any constraint, the loss-averse investor follows a partial portfolio insurance strategy: in good states the investor tries to insure the reference point of the utility, while in bad states the allocation to stocks is increased due to the risk-seeking behavior. Therefore, implementing a portfolio insurance strategy for a loss averse investor with the optimal allocation program could potentially lead to results different from those obtained in the studies using constant multipliers mentioned in the previous paragraph. In particular, this raises the question of the usefulness of portfolio insurance for loss-averse investors implementing an optimal strategy.

Several authors have examined optimal portfolios with capital protection. In particular, Cox and Leland (1982) and Grossman and Vila (1989) study the portfolio optimization problem when a constraint is imposed on the terminal date, while El Karoui, Jeanblanc, and Lacoste (2005) and El Karoui and Meziou (2006) also consider a constraint imposed on every intermediary date. Basak and Shapiro (2001) derive the optimal portfolio for an investor with a constraint expressed in terms of the probability of the wealth to fall below some floor, while Boyle and Tian (2007) consider an investor desiring to outperform some benchmark with a certain confidence level at maturity. Bouchard, Elie, and Touzi (2009) and Bouchard, Elie, and Imbert (2010) also study the problem associated which guarantees to reach a controlled target with a given probability. In this paper, we study the characteristics and attractiveness of optimal investment strategies with capital protection expressed in terms of the running maximum of wealth. Namely, we consider the widely used drawdown constraint that the wealth process never falls below a fixed fraction of its running maximum. The optimal investment problem of this type of strategy for an agent maximizing the long term growth rate of the expected utility of final wealth has been studied by Grossman and Zhou (1993), while Cvitanic and Karatzas (1995) consider a more general class of price generating processes. Cherny and Obłój (2013) also study a semimartingale financial market model. Furthermore, Elie and Touzi (2008) derive the optimal investment strategy corresponding to the infinite horizon optimal consumption problem under the drawdown constraint, and Elie (2008) considers the finite time horizon case. The results of the aforementioned studies show that the optimal investment in the risky asset is a function of the difference between the current wealth and its running maximum, the risk aversion of the investor, and the parameters of the risky asset generating process.

Based on a large panel of numerical experiments produced from the resolution of the stochastic control problem corresponding to the optimal investment problem under drawdown constraint, we compute numerically and analyze optimal allocation programs (i.e., optimal multipliers). Our numerical approach allows to study a general class of utility functions. We consider a S-shape utility (proposed in the prospect theory of Kahneman

and Tversky, 1979; Tversky and Kahneman, 1992) differing between the domains of gains and losses with respect to the reference point, and a standard concave utility (Constant Relative Risk Aversion or power utility). The finite horizon expectation maximization problem is studied and we highlight the divergences in investment performance and optimal behavior of investors according to their utility function, the parameters of the risky asset price process, and the presence or absence of the drawdown constraint. We compute the optimal investments strategies, by solving the Hamilton–Jacobi–Bellman equation, that characterizes the dynamic programming principle related to the stochastic control problem, making vary the market model parameters and utility profiles. The resulting Partial Differential Equation (PDE) problem is then (numerically) solved on a subset of $\mathbb{R}^2$ with a boundary condition that characterize the drawdown constraint.

Using a stochastic sampling of the market, we find that, from an optimal standpoint, a general pattern justifying the potential attractiveness and usefulness of strategies with drawdown control cannot be drawn in terms of Certainty Equivalent and terminal utility maximization, in a prospect theory context. First, for a CRRA investor with power utility, the preference for the optimal strategy with drawdown constraint rather than the optimal strategy without protection, strongly depends on the values of the parameters of the stock market dynamics. On the other hand, in our simulations, we find that the prospect theory investor with S-shape utility function rather prefers the optimal strategy without drawdown constraint. This result can be partially explained by the presence of the reference point in the S-shape utility function. Indeed, as discussed by Berkelaar et al. (2004), optimal portfolios without any constraint for the latter utility reveal that this reference point acts as an "implicit protection", implying that the drawdown limitation is not necessarily useful for this type of investor if the optimal strategy is implemented. In other words, the optimal strategy with drawdown constraint does not lead to larger final utility values with respect to the optimal strategy without constraint. These results hold even when we consider a short investment horizon of one year. Nevertheless, the only setups where the drawdown constraint could be gainful for this investor correspond to the cases where the model's parameters are mis-estimated. Indeed, overestimation of the risk premium when the volatility is at a substantial level, and severe underestimation of the volatility when the risk premium is low, can lead to larger terminal utility in the presence of the drawdown constraint. In these specific cases, the drawdown limitation can act as an insurance against model's mis-estimation risk.

The paper is organized as follows. Section 3.2 introduces the model for the stock market. Section 3.3 is devoted to the examination of the drawdown constraint and the optimization problem. Section 3.4 and 3.5 present the methodology we employ in order to derive the optimal control. Our numerical resolution procedure is described in section 3.6. Finally our simulations and main results are discussed in section 3.7.

3.2 The model

We consider a complete filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$, where the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ is generated by the standard Brownian motion Z. In the financial market, investors can trade two assets, S_t and S_t^0, respectively a risky asset and a risk-less money account with general dynamics:

$$dS_t = \mu(t)S_t dt + \sigma(t)S_t dZ_t, \tag{3.1}$$

and

$$dS_t^0 = r(t)S_t^0 dt, \tag{3.2}$$

where the drift μ, volatility σ and interest rate r, are bounded adapted processes. We also denote by

$$\beta_t := \exp\left(-\int_0^t r(s)ds\right),\ t \geq 0$$

the discount process of the model.

Recent papers studying the popularity of portfolio insurance[2] suggest that the demand for such strategies could be justified in a behavioral finance context and not using standard utility theory. Therefore, we consider two different types of investors through their utility function, namely 1) a S-shape utility (prospect theory), and 2) a concave power utility, i.e., Constant Relative Risk Aversion (CRRA). The latter utility is usually considered as a benchmark as it is the most frequently assumed specification. Following the results of Kahneman and Tversky (1979), we consider a loss-averse prospect theory investor with the following asymmetric utility function over gains and losses relative to its reference point Θ:

$$U(x,\Theta) = \begin{cases} -A(\Theta - x)^{\gamma_1} & \text{for } x \leq \Theta, \\ +B(x-\Theta)^{\gamma_2} & \text{for } x > \Theta, \end{cases} \tag{3.3}$$

where A, B, γ_1 and γ_2 are constant parameters. Through our study, we consider a *discounted reference level* Θ constant over time. Subjective probability distortion is also an important aspect of prospect theory (see for instance Lattimore, Baker, and Witte, 1992). In this paper, we do not use subjective probability weights as our objective is to study the impact of the choice of utility on the attractiveness of the drawdown constraint.

The standard CRRA utility we consider is given by:

$$U(x,\Theta) = U(x) = \frac{x^{1-\gamma}}{1-\gamma} \quad \text{for } x > 0, \tag{3.4}$$

where γ is constant. Note that the theoretical reference level Θ in equation (3.4) is only presented for the sake of uniformity of the function's parametrization between the S-shape and the CRRA utility. The CRRA investor has no interest in an hypothetical reference level. Figure 3.1 exhibits the two utility functions with the parameter set we use in our

[2] cf., Vrecko and Branger (2009); Dierkes et al. (2010); Dichtl and Drobetz (2011); Ebert et al. (2012); Khuman et al. (2012).

empirical procedure of section 3.7.

3.3 Portfolio management with drawdown constraint

We consider the following optimal investment problem, with capital protection expressed in terms of running maximum, described by an investor *discounted wealth process* $W^{x,\pi}$ satisfying

$$\begin{aligned} dW_t^{x,\pi} &= \pi_t \left(W_t^{x,\pi} - \alpha M_t^{x,\pi}\right) \left\{(\mu(t) - r(t))dt + \sigma(t)dZ_t\right\}, \\ W_0^{x,\pi} &= x, \end{aligned} \tag{3.5}$$

with a given protection level $\alpha \in (0,1)$ and the *discounted wealth running maximum* denoted

$$M_t^{x,\pi} = \sup_{0 \le s \le t} W_s^{x,\pi}.$$

π is the multiplier of the portfolio insurance strategy, also denoted hereafter as the *control*. At any time, the investor makes her investment decision proportional to the amount $(W_t^{x,\pi} - \alpha M_t^{x,\pi})$. As a consequence the discounted wealth satisfies the strict drawdown constraint

$$W_t^{x,\pi} > \alpha M_t^{x,\pi}, \text{ for all } 0 \le t \le T, a.s., \tag{3.6}$$

where T stands for the finite investment horizon of interest.

The optimization problem we consider is the following:

$$\text{for } 0 < T < \infty, \text{ find } \pi^* = \arg\max_{\pi \in \mathcal{A}} \mathbb{E}\left[U(W_T^{x,\pi}, \Theta)\right], \tag{3.7}$$

for a set $\mathcal{A}$ of admissible investment strategies π adapted to the filtration $(\mathcal{F}_t, t \ge 0)$ generated by the Brownian motion Z, and valued in $[\pi_{\min}, \pi_{\max}]$ with $\pi_{\min}, \pi_{\max}$ finite[3]. The range $[\pi_{\min}, \pi_{\max}]$ can be determined according to eventual allocation constraints (e.g., no short-selling and maximal leverage).

We recall that $M^{x,\pi}$ is non-decreasing and varies only when it equals $W^{x,\pi}$, that is

$$\int_0^T (M_t^{x,\pi} - W_t^{x,\pi}) dM_t^{x,\pi} = 0. \tag{3.8}$$

We also recall some mathematical consequences of our model formulation. The expression of the discounted value of wealth $W^{x,\pi}$ with $0 < \alpha < 1$ can be derived as a "corrected" version of the solution without constraint. Indeed, when $\alpha = 0$, for a given admissible strategy $\pi \in \mathcal{A}$, equation (3.5) has a well known solution form

$$w_t^{x,\pi} = x \exp\left(\int_0^t \left\{\pi_s \left(\mu(s) - r(s)\right) - \frac{1}{2}\sigma(s)^2 \pi_s^2\right\} ds + \int_0^t \pi_s \sigma(s) dZ_s\right) \tag{3.9}$$

[3]Cvitanic and Karatzas (1995) and Elie and Touzi (2008) provide a formal definition of the set of admissible strategies when the control is unbounded.

and we denote

$$m_t^{x,\pi} = \sup_{0\leq s\leq t} w_s^{x,\pi}.$$

The following proposition gives the expression of the general solution.

Proposition 6 *For all admissible strategies $\pi \in \mathcal{A}$, the discounted wealth equation* (3.5) *has a unique explicit solution*

$$W_t^{x,\pi} = \left[w_t^{x,\pi} + \frac{\alpha}{1-\alpha} m_t^{x,\pi}\right] \left(\frac{m_t^{x,\pi}}{1-\alpha}\right)^{-\alpha} \tag{3.10}$$

$$M_t^{x,\pi} = \left(\frac{m_t^{x,\pi}}{1-\alpha}\right)^{(1-\alpha)} \tag{3.11}$$

and the strict drawdown constraint (3.6) *is always satisfied.*

We refer to Cvitanic and Karatzas (1995) for the first derivation of this expression and to Elie (2006) and Elie and Touzi (2008) for a formal proof.

3.4 The Hamilton-Jacobi-Bellman equation

We shall use the dynamic programming approach in order to derive the Partial Differential Equation (PDE) formulation of the optimization problem (3.7). We then need to introduce a dynamic version of this problem within the value function:

$$v(t,x,y) = \sup_{\pi\in\mathcal{A}_t} \mathbb{E}_{x,y}\left[U(W_T^{\pi}, \Theta)\right], \tag{3.12}$$

defined for the pair (x,y), with the restriction $0 < \alpha y < x \leq y$, as the initial condition of the state processes $(W^{x,\pi}, M^{x,\pi})$ starting at time $0 \leq t < T$ with dynamics

$$\begin{cases} W_\tau^{x,\pi} = x + \displaystyle\int_t^\tau \pi_s \left(W_s^{x,\pi} - \alpha M_s^{x,\pi}\right)\left\{(\mu(s) - r(s))ds + \sigma(s)dZ_s\right\}, \\ M_\tau^{x,\pi} = y \vee \displaystyle\sup_{t\leq s\leq\tau} W_s^{x,\pi}. \end{cases}$$

$\mathcal{A}_t$ is the set of admissible strategies on the sub-period $[t,T]$ defined as the subset of $\mathcal{A}$ of strategy π independent of $\mathcal{F}_t$.

The value function $v(t,x,y)$ can be easily shown to be non decreasing with respect to x and concave when the utility function U is. Moreover for any $y \geq 0$, for any $0 \leq t \leq T$, $v(t,\alpha y, y) = U(\alpha y, \Theta)$, i.e., for all $0 \leq t \leq T, W_t^{\alpha y,\pi} = \alpha y$. In other words, if the initial wealth is equal to αy the wealth does not change over time. In particular, following Elie and Touzi (2008), applying at least formally the Itô formula (assuming the required

regularity on v), we get

$$\begin{aligned}
&\sup_{\pi\in\mathcal{A}_t}\mathbb{E}_{x,y}[v(t+h,W_{t+h},M_{t+h})]\\
&=v(t,x,y)+\sup_{\pi\in\mathcal{A}_t}\mathbb{E}_{x,y}\int_t^{t+h}\frac{\partial v}{\partial x}(s,W_s,M_s)dW_s \qquad (3.13)\\
&\quad+\sup_{\pi\in\mathcal{A}_t}\mathbb{E}_{x,y}\int_t^{t+h}\left\{\frac{\partial v}{\partial s}(s,W_s,M_s)+\mathcal{L}_{\pi_s}(s)v(s,W_s,M_s)\right\}ds \qquad (3.14)\\
&\quad+\sup_{\pi\in\mathcal{A}_t}\mathbb{E}_{x,y}\int_t^{t+h}\frac{\partial v}{\partial y}(s,W_s,M_s)dM_s \qquad (3.15)
\end{aligned}$$

with the second-order operator

$$\mathcal{L}_p(t)\phi=p(x-\alpha y)(\mu(t)-r(t))\frac{\partial\phi}{\partial x}+p^2(x-\alpha y)^2\frac{\sigma^2(t)}{2}\frac{\partial^2\phi}{\partial x^2}, \qquad (3.16)$$

where $\phi\in\mathcal{C}^2(\mathbb{R})$.

Using the dynamic programming principle we can derive the associated Hamilton–Jacobi-Bellman equation. First, we use the fact that the expectation of the stochastic integral in (3.13) vanishes under the formal a priori regularity hypothesis on the value function. Then, dividing by h and taking the limit as $h\longrightarrow 0$, the guessed associated Hamilton–Jacobi–Bellman equation for $v(t,x,y)$ is:

$$\frac{\partial v}{\partial t}+\max_{p\in[\pi_{\min},\pi_{\max}]}\left[p(x-\alpha y)(\mu(t)-r(t))\frac{\partial v}{\partial x}+p^2(x-\alpha y)^2\frac{\sigma^2(t)}{2}\frac{\partial^2 v}{\partial x^2}\right]=0,\ \text{for } x<y, \qquad (3.17)$$

where we impose the natural condition $\dfrac{\partial v}{\partial y}(t,x,y)=0,\text{on }\{x=y\}$, in order to equal to zero the expectation term in (3.15), using the fact that the maximum only varies when $x=y$, cf., equation (3.8).

Therefore, the boundary conditions considered are the following:

$$\begin{cases}
v(T,x,y)=U(x,\Theta), & \text{for } \alpha y<x\leq y,\\
v(t,\alpha y,y)=U(\alpha y,\Theta) & \text{for } 0\leq t\leq T\\
\dfrac{\partial v}{\partial y}(t,x,y)=0, & \text{on } \{x=y\}.
\end{cases}$$

Note that when $\alpha=0$ we end up with Merton (1969)'s problem (see for instance Duffie, 2010, on the theory of portfolio choice and asset pricing, including discussion of Merton's work and next contributions).

On an infinite time horizon, Elie (2006) and Elie and Touzi (2008) present a rigorous mathematical proof that for a concave utility function U, the required regularity conditions on the value function are satisfied in order to derive a unicity and verification theorem[4]. Elie (2008) considers the finite time horizon case. For the S-shape utility, the derivation

[4]Boyle and Tian (2007) also derive a rigorous mathematical analysis of this kind of problem for an investor who desires to outperform some benchmark index with a certain confidence level.

of a similar result constitutes an entire problem beyond the scope of this paper. Our main analysis is focused on the study of strategies implementing the optimal control in stochastic simulations.

3.5 Heuristic for the optimal control study

3.5.1 Pontryagin maximum principle

Based on verification conjecture, in the following calculation we derive the feedback control candidate $p^*(t,x,y)$ using the Pontryagin maximum principle. The Hamiltonian corresponding to the optimization problem is defined as:

$$\mathcal{H}\left(p;t,x,y,v,\frac{\partial v}{\partial x},\frac{\partial^2 v}{\partial x^2}\right)=p(x-\alpha y)(\mu(t)-r(t))\frac{\partial v}{\partial x}+p^2(x-\alpha y)^2\frac{\sigma^2(t)}{2}\frac{\partial^2 v}{\partial x^2}. \quad (3.18)$$

The feedback control candidate $p^*(t,x,y)$ is identified as

$$p^*(t,x,y)=\underset{p\in[\pi_{\min},\pi_{\max}]}{\arg\max}\ \mathcal{H}\left(p;t,x,y,v,\frac{\partial v}{\partial x},\frac{\partial^2 v}{\partial x^2}\right). \quad (3.19)$$

For any (t,x,y), we denote

$$p_0(t,x,y):=\frac{1}{\sigma(t)^2(x-\alpha y)}\times\frac{-(\mu(t)-r(t))\dfrac{\partial v}{\partial x}}{\dfrac{\partial^2 v}{\partial x^2}},$$

the control value resulting from the first order condition applied to the Hamiltonian. The second order derivative of the Hamiltonian $\mathcal{H}$ with respect to p is equal to

$$\frac{\partial^2\mathcal{H}}{\partial p^2}=\sigma^2(x-\alpha y)^2\frac{\partial^2 v}{\partial x^2}.$$

Hence, the value $p^*(t,x,y)$ of the optimal control at the state (t,x,y) depends on the sign of $\frac{\partial^2 v}{\partial x^2}$ in the following way

$$p^*(t,x,y):=\begin{cases} p_0(t,x,y) & \text{if } \frac{\partial^2 v}{\partial x^2}<0 \text{ and } p_0(t,x,y)\in[\pi_{\min},\pi_{\max}],\\ \underset{\pi\in\{\pi_{\min},\pi_{\max}\}}{\arg\max}\ \mathcal{H}\left(\pi;t,x,y,v,\frac{\partial v}{\partial x},\frac{\partial^2 v}{\partial x^2}\right) & \text{otherwise.}\end{cases} \quad (3.20)$$

Hence p^* is the unique solution of the quadratic problem given by equation (3.19).

3.5.2 Intuition on the allocation profiles

Before considering the numerical resolution of the problem in the next section, we derive below a first intuition on the behavior of the different types of investor.

We consider first the prospect theory investor with utility function given by (3.3). If

$x \leq \Theta$, we have:

$$\frac{\partial^2 U}{\partial x^2}(x,\Theta) = -A\gamma_1(\gamma_1 - 1)(\Theta - x)^{\gamma_1 - 2}.$$

If $A > 0$ and $0 < \gamma_1 < 1$ as in our main analysis, the second derivative of the utility with respect to x is thus positive. Therefore, assuming $v(t,x,y) \approx U(x,\Theta)$ for any $0 \leq t \leq T$, for illustration purpose only, the value of p^* will not be equal to p_0. Given the positive sign of the second derivative above and the expression of the Hamiltonian, the value of p^* will be $\pi_{\max}$. Hence according to this approximation for the prospect theory investor, if the wealth is below the reference point, the optimal control is likely to be the maximal value $\pi_{\max}$, emphasizing the risk-seeking behavior.

On the other hand, if $x > \Theta$, we get:

$$\frac{\partial^2 U}{\partial x^2}(x,\Theta) = B\gamma_2(\gamma_2 - 1)(x - \Theta)^{\gamma_2 - 2}.$$

If $B > 0$ and $0 < \gamma_2 < 1$ as in our main analysis, the second derivative of the utility with respect to x is thus negative. Therefore, assuming similarly $v(t,x,y) \approx U(x,\Theta)$ for any $0 \leq t \leq T$, for illustration purpose only, the value of p^* can be equal to p_0 given that $p_0 \in [\pi_{\min}, \pi_{\max}]$. In that case, the value of p^* is given by

$$p_0(t,x,y) = -\frac{\mu - r}{\sigma^2(x - \alpha y)} \times \frac{(x - \Theta)}{\gamma_2 - 1}.$$

This expression clearly highlights that as $x \downarrow \Theta$, the control tends to zero, i.e., as the wealth of the prospect theory investor declines and gets closer to the reference point, the investor is likely to reduce the allocation to the risky asset in order to ensure the amount Θ. Note that this approximation is valid whether or not the drawdown constraint is considered, i.e., even if $\alpha = 0$.

Finally these approximations are in line with Berkelaar et al. (2004)'s results (related to the case without drawdown constraint). Indeed, for the latter configuration, Berkelaar et al. (2004) show that the optimal strategy is a partial portfolio insurance: in good states the investor tries to insure the reference point, while in bad states the investor increases the allocation to stocks due to the risk-seeking behavior.

For the investor with power utility function given by (3.4), the result is different. We get:

$$\frac{\partial^2 U}{\partial x^2}(x,\Theta) = -\gamma x^{-\gamma - 1}.$$

Hence the second derivative of the utility with respect to x is negative. Therefore, assuming again for illustration purpose that $v(t,x,y) \approx U(x,\Theta)$ for any $0 \leq t \leq T$, the value of p^* can be equal to p_0. Given that $[\pi_{\min}, \pi_{\max}]$, we obtain that p^* is equal to:

$$p_0(t,x,y) = \frac{\mu - r}{\sigma^2(x - \alpha y)} \times \frac{x}{\gamma}.$$

As expected, this expression should imply a very different behavior for the CRRA investor given the absence of reference point.

3.6 Numerical resolution

In this section, we present our approach for the resolution of the Hamilton-Jacobi-Bellman equation. We aim to consider a general class of utility functions, even those for which a closed form solution to the problem (3.7) does not exist. Thus, in order to perform an extensive analysis of optimal investments under the drawdown constraint, we solve numerically the optimization problem[5]. For that purpose, we have to limit the space considered, i.e., we define a discrete space for the values of the wealth and the running maximum where the optimization problem will be solved.

First we define a discretization of the time period of interest $[0, T]$: $\{t_k = k\Delta t, k = 0, ..., K\}$, where Δt is the discretization time step and $t_K = K\Delta t = T$. Secondly, for each time step, we define a 2D grid $G = \{(x_i, y_j), i = 0, ..., N_x; j = 0, ..., N_y\}$ on the space $(W^{x,\pi}, M^{x,\pi})$, with steps Δx and Δy, i.e., $x_i = x_0 + i\Delta x$ and $y_j = y_0 + j\Delta y$. x_0 and x_{N_x} are respectively the minimal and maximal value of wealth considered. Similarly $y_0 \geq x_0$ and y_{N_y} are respectively the minimal and maximal value of running maximum considered. Given that we normalize the initial value of wealth to one in our empirical procedure, we set $x_0 = \alpha$ and $y_0 = 1$. The maximal values x_{N_x} and y_{N_y} are defined according to the parameters of the risky asset dynamics in our simulations, discussed in section 3.7, such that the occurrence of values of wealth and running maximum above x_{N_x} and y_{N_y} is very unlikely. Moreover, given the high values of x_{N_x} and y_{N_y} used in our analysis, we add the additional artificial boundary condition $\frac{\partial v}{\partial x} = 0$, on $x = x_{N_x}$ and $y = y_{N_y}$. Therefore, for $x = x_{N_x}$ and $y = y_{N_y}$ (and values beyond the boundaries) the optimal control will be equal to zero given equation (3.20), which is intuitive given the flatness of the utility functions for these levels of wealth.

In order to derive an approximation of the optimal control p^* for each point on the grid, we use the finite difference method, applied to the non linear operator $\mathcal{L}_p$. For a given time step t_k, let $v^k_{i,j}$ be a discrete approximation of $v(t_k, x_i, y_j)$ and let $\mathcal{L}^k_p$ be the discretized version of the differential operator $\mathcal{L}_p$ defined by (3.16), on the grid G. This operator can be discretized using spacial forward, centered or backward differencing in the x and y directions to give:

$$\begin{aligned}(\mathcal{L}^k_p v^k)_{i,j} = {} & \alpha^k_{i,j}(p)v^k_{i-1,j} + \beta^k_{i,j}(p)v^k_{i+1,j} + \gamma^k_{i,j}(p)v^k_{i,j-1} + \delta^k_{i,j}(p)v^k_{i,j+1} \\ & - (\alpha^k_{i,j}(p) + \beta^k_{i,j}(p) + \gamma^k_{i,j}(p) + \delta^k_{i,j}(p))v^k_{i,j}, \end{aligned} \tag{3.21}$$

where p stands for a given value of the control and $\alpha^k_{i,j}, \beta^k_{i,j}, \gamma^k_{i,j}$ and $\delta^k_{i,j}$ are coefficient functions depending on the value of the control and the model parameters. Note that $\gamma^k_{i,j}(p) = \delta^k_{i,j}(p) = 0$ for all i, j and p, given the absence of derivatives with respect to y in (3.16). Therefore equation (3.21) reduces to:

$$(\mathcal{L}^k_p v^k)_{i,j} = \alpha^k_{i,j}(p)v^k_{i-1,j} + \beta^k_{i,j}(p)v^k_{i+1,j} - (\alpha^k_{i,j}(p) + \beta^k_{i,j}(p))v^k_{i,j}. \tag{3.22}$$

Following Forsyth and Labahn (2008), the discretization method must maintain the mono-

[5] Elie (2008) also considers a numerical resolution of the HJB equation associated with the drawdown constraint.

tonicity of $\mathcal{L}_p^k$ applied to v^k in order to insure the consistency of our approach. In Appendix 3.A, we derive the conditions on the differencing procedure of the value function in line with this requirement.

At each time step, we have to solve the HJB equation given by equation (3.17) in order to derive the value of p^*. We use the following discretization based on an implicit scheme in order to ensure the stability of the algorithm[6] for any given value of Δt:

$$\frac{v_{i,j}^{k+1} - v_{i,j}^{k}}{\Delta t} + \max_{\pi_{k+1} \in \mathcal{A}} (\mathcal{L}_{\pi_{k+1}}^{k+1} v^{k+1})_{i,j} = 0, \tag{3.23}$$

where i, j indicate the point (x_i, y_j) on the grid G. The iteration policy is the following. For a given time step t_{k+1}, first we use the value of the optimal control calculated for the previous time step as a first guessed candidate $\left\{p_{k+1}^*\right\}_0$. Using this value of the control, we derive the corresponding initial approximation of the value function $\left\{\hat{v}_{i,j}^{k+1}\right\}_0$ for the time step t_{k+1} using the discrete approximation of the derivatives of the value function appearing in the operator $\mathcal{L}_p$. Then we build the iteration matrix $\left\{\left(\mathcal{L}_{p^*}^{k+1}\hat{v}^{k+1}\right)_{i,j}\right\}_0$ for the grid G. We solve again the HJB equation (3.23) and obtain a new value function $\left\{\hat{v}_{i,j}^{k+1}\right\}_1$ and the corresponding approximation of the optimal control $\left\{p_{k+1}^*\right\}_1$. We repeat the process until we reach an iteration $n+1$ where we observe that

$$\frac{\left|\left\{\hat{v}_{i,j}^{k+1}\right\}_{n+1} - \left\{\hat{v}_{i,j}^{k+1}\right\}_n\right|}{\left|\left\{\hat{v}_{i,j}^{k+1}\right\}_{n+1}\right|} < \Lambda,$$

where Λ is the tolerance value (equal to 10^{-6} in our numerical procedure), i.e., we consider that the value function has converged. In that case, the corresponding current optimal control value $\left\{p_{k+1}^*\right\}_{n+1}$ is extracted and used as an approximation of p^* for the corresponding time step t_{k+1}.

3.7 Empirical analysis based on simulated markets

3.7.1 Stochastic sampling of the markets and portfolios

The objective of our study is to analyze the numerical simulations and resolutions of the model. For that purpose, we perform simulations of the risky asset according to the dynamics given in equation (3.1). In our simulations, parameters μ, r, σ are assumed to be constant on the period $[0, T]$. Therefore, the solution of equation (3.1) is given by:

$$S_t = S_0 \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right). \tag{3.24}$$

Given Δt the discretization time step and discrete times $t_k = k\Delta t, 0 \leq k \leq K$, the

[6] In particular the implicit scheme allows us to choose the time step Δt independently of the space step.

risky asset sample paths are generated as follows:

$$S_{t_{k+1}} = S_{t_k} \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right)\Delta t + \sigma\sqrt{\Delta t} Z_{t_{k+1}}\right), \tag{3.25}$$

where $S_0 = 1$, and $\{Z_{t_{k+1}}, 0 \leq k \leq N-1\}$ are independent realizations of a standard Gaussian random variable.

For each sample path, we simulate exactly the wealth corresponding to the drawdown control strategy and its running maximum (both with initial value equal to 1) using the approximated control p^* derived using the methodology described in the previous section and collected on the grid G. Namely, at each time step, the level of wealth and running maximum are simulated exactly (cf., below), and according to these current levels (of wealth and running maximum) we find the closest point (x_i, y_j) on the grid G, and implement the corresponding value of the optimal control p^* in the strategy. Note that the derivatives of the value function appearing in equation (3.20) are evaluated using the procedure described in Appendix 3.A. Portfolio dynamics presented in the paper correspond to discounted values. Hence all amounts are simulated and evaluated in terms of their discounted values.

From Proposition 6, the simulation of $(W_t^{1,p^*}, M_t^{1,p^*}, t \in [0,T])$ reduces to the simulation of $(w_t^{1,p^*}, m_t^{1,p^*}, t \in [0,T])$ solution of (3.5). We emphasize that, based on the assumption that the control is constant between each time step, the simulation method we use for $(w^\pi, m^\pi, W^\pi, M^\pi)$ is exact. Let $\{\pi_k, k = 0, ..., K\}$ be the values of an admissible control at each time step. Assuming the portfolio rebalancing strategy π is constant on $[t_k, t_{k+1}[, k = 0, ..., K-1$, one has to simulate exactly the random variable:

$$w_{t_{k+1}}^\pi = w_{t_k}^\pi \exp\left(\left[\pi_k(\mu - r) - \frac{1}{2}\sigma^2\pi_k^2\right]\Delta t + \pi_k\sigma\sqrt{\Delta t} Z_{t_{k+1}}\right), \tag{3.26}$$

and its running maximum $m_{t_{k+1}}^\pi$, in order to construct the trajectory of $W_{t_k}^\pi$ and $M_{t_k}^\pi$, the discrete versions of the discounted wealth process $W_{t_k}^{1,\pi}$ and running maximum $M_{t_k}^{1,\pi}$ respectively, as:

$$W_{t_k}^\pi = \left[w_{t_k}^\pi + \frac{\alpha}{1-\alpha} m_{t_k}^\pi\right]\left(\frac{m_{t_k}^\pi}{1-\alpha}\right)^{-\alpha}, \tag{3.27}$$

$$M_{t_k}^\pi = \left(\frac{m_{t_k}^\pi}{1-\alpha}\right)^{(1-\alpha)}. \tag{3.28}$$

We provide further details on the exact simulation of the joint dynamics of w_t^π and m_t^π in Appendix 3.B.

The results of the numerical simulations and resolutions of the model are presented for different parameter sets and utility functions. In our initial set of simulations, we consider an investment horizon of five years with a daily time step, i.e., $\Delta t = 1/252$, and $K = 5 \times 252$. The initial value of wealth is arbitrary set to 1 in all our simulations. The space steps $(\Delta x, \Delta y)$ for the resolution of the control problem are set to 1.0%, the minimal values of wealth and running maximum on the grid are set to α and 1, while the

maximal values we consider on the grid are both equal to 7.5. The allocation programs correspond to a procedure with continuous rebalancing, i.e., daily allocation adjustment based on available information. In a second stage, we also consider a shorter horizon of one year, i.e., $K = 252$, with daily time step.

The protection level α is set to 0.8, which is in line with the drawdown level quoted by Grossman and Zhou (1993)[7]. In addition, we also consider strategies without capital protection, that is $\alpha = 0$. As mentioned above, we use constant parameters for the dynamics of assets. Without loss of generality r is arbitrarily set to 0.03, and we consider several parameters for the risky asset dynamics: $\mu \in \{0.10, 0.05\}$, i.e., high and low risk premium, and $\sigma \in \{0.10, 0.20, 0.30\}$, i.e., low, medium and high volatility, which leads to 6 market configurations, summarized in Table 3.1. Note that the maximal values of 7.5 for wealth and running maximum on our grid are coherent with these market dynamics. For each market configuration, we simulate 50,000 sample paths for the risky asset and run the different portfolio strategies for the different investors described below.

As mentioned in section 3.2, we consider two types of investors. The first one has a prospect theory S-shape utility function described by equation (3.3). The parameter set of the latter is borrowed from Kahneman and Tversky (1979), and is defined as follows: $A = 2.25, B = 1$, and $\gamma_1 = \gamma_2 = 0.88$. The *discounted reference level* is constant and set to the initial wealth for this investor, i.e., $\Theta = 1$. We also consider additional resolutions and simulations with $\Theta = 1.05$, i.e., a reference point higher than the initial level of wealth. The second type of investors has a CRRA utility function described by equation (3.4). The risk aversion coefficient γ is set to 2, as it represents the usual benchmark utility function.

For each investor, capital protection, and asset parameters, the discrete optimal control is derived using the method described in sections 3.5 and 3.6. The admissible interval $[\pi_{\min}, \pi_{\max}]$ is defined according to allocation constraints. Indeed, we allow the optimal strategies to invest between 0% and 200% in the risky asset. Therefore, for each configuration, we set $\pi_{\min} = 0$, and $\pi_{\max} = 2/(1 - \alpha)$. This choice leads to $\pi_{\max} = 2$ without capital protection ($\alpha = 0$), and $\pi_{\max} = 10$ when $\alpha = 0.8$. The results are discussed below. We also consider constant control as benchmark investment strategies.

We present tables containing several performance metrics. We focus especially on Certainty Equivalent (CE)[8], and average final utility across simulations: the main portfolio selection criteria. Statistics related to annualized rate of return, volatility, skewness and kurtosis, maximum cumulative loss (drawdown), Sharpe and Calmar[9] ratios are examined as well, in terms of average value across scenarios. The average allocation to stocks and the average control over each trajectory and over all sample paths are also presented.

[7] In our numerical procedure, we also considered a protection level α equal to 0.5. The overall qualitative results do not change. Corresponding tables are presented in Appendix 3.C.

[8] Certainty Equivalent is defined as the deterministic level of terminal wealth that gives the investor the same expected utility as the terminal wealth resulting from the strategy.

[9] The Calmar ratio is calculated as the annualized return divided by the maximum drawdown of the portfolio trajectory.

3.7.2 Empirical results

In this section, we discuss the numerical results of our simulations for different asset generating processes, investors and protection levels. Except for section 3.7.3, the reference point Θ of the S-shape utility is set to the initial level of wealth, i.e., one.

Strategies implementing a constant control

First, as benchmark strategies without any optimization procedure, we simulate portfolios with constant control over five years. We consider strategies without capital protection, i.e., $\alpha = 0$, and with capital protection $\alpha = 0.8$. For the case $\alpha = 0$, we implement first a constant control equal to 0.5, this configuration corresponds to a pure constant mix strategy, i.e., a combination of 50% of the risky asset and 50% of the riskless asset. We also consider a constant control equal to 1, corresponding to a pure investment in the risky asset. For the case, $\alpha = 0.8$, we implement two different sets of strategies with two constant controls equal to 5 and 10. The ex-post statistics in terms of risks and returns are presented in Tables 3.2 and 3.3 respectively for the simulations without and with drawdown constraint. Note that these metrics are common to both investors, as the wealth trajectories are identical. Indeed due to the fact that the constant control is independent of the utility function, the simulations lead to identical results. Overall, a clear observation is the better performance in terms of average (annualized) return and Sharpe ratio of the strategy with loss control when the risk premium is low (right part of the tables) for all values of volatility (note however that the strategy fully invested in risky asset seems to be the best one when the volatility is low). The usefulness of the protection can be also highlighted with the minimal annualized returns over all scenarios, which are much lower in the absence of the constraint. On the other hand, when the risk premium is high (left part of the tables), the protected strategy leads to larger mean return only in the most favorable configuration (lowest volatility), with respect to the constant control 0.5. In the two remaining setups the opportunity cost of the drawdown constraint appears to penalize the strategy with respect to the unprotected portfolio.

Regarding the preference for one type of strategy rather than the other, we present the results concerning the utility measurements for the CRRA investor in Tables 3.4 and 3.5. We can observe that, in terms of Constant Equivalent, the CRRA 2 investor does not systematically prefer one type of strategy with or without drawdown constraint. As suggested by the return figures, on the one hand, when the risk premium is low ($\mu - r = 0.02$), the investor always prefers the capital protection. Indeed, the Certainty Equivalent and average terminal utility are greater for the strategies with $\alpha = 0.8$ (with control equal to 5 and 10) than for the strategies without protection. This preference can be justified by better (risk-adjusted) performance (discussed above) and lower risk (lower mean annualized volatility and maximum drawdown except when $\sigma = 0.1$). However, these observations do not hold when the risk premium is higher. While both strategies with capital protection lead to a higher CE for the configuration $\mu - r = 0.1, \sigma = 0.1$ with respect to the case $\alpha = 0, \pi = 0.5$, they are dominated by both strategies with $\alpha = 0$

when $\sigma = 0.2$, and by the portfolios $\pi = 0.5$ for the remaining configuration ($\sigma = 0.3$). Indeed, we can see in Tables 3.2 and 3.3 that the capital protection implies a smaller allocation to the risky asset, leading to lower annualized returns. Therefore, according to our simulations with constant control, the preference for capital protection or not remains dependent on market conditions for the CRRA investor with power utility.

We examine now the corresponding results obtained for the investor with S-shape utility, presented in Tables 3.6 and 3.7. In fact, the conclusions remain quite similar in terms of ex-post utility and Constant Equivalent as well. Again, when the risk premium is low, the CE obtained for the prospect theory investor with protection are greater than those with $\alpha = 0$ (except when $\sigma = 0.1$: better performance of the portfolios with $\pi = 1$). When the risk premium is high, we obtain the same preference for the strategy with $\alpha = 0$ when $\sigma = 0.2$ and $\sigma = 0.3$. Finally, the results also strongly depend on the parameters of the risky asset dynamics. Note that as expected, for both investors, average final utility and Constant Equivalent (CE) decrease as $\mu - r$ decreases and as σ increases.

Strategies implementing the optimal control

The results obtained with the strategies using the optimal control p^* derived in section 3.5 over five years are now discussed.

As illustrated in Table 3.8, for the CRRA 2 investor, the strategies without capital protection implementing the optimal control lead to better results than similar strategies with constant control equal to 0.5 and 1 discussed above. Indeed, in all market configurations, the strategies using p^* imply greater values in terms of Constant Equivalent/terminal utility (and Sharpe ratio). We note that the average "optimal" allocations to the risky asset differ clearly from 50% in each configuration, and range from 200% (when $\mu - r = 0.07$ and $\sigma = 0.1$) to only 11% (when $\mu - r = 0.02$ and $\sigma = 0.3$). In addition, we observe that the magnitude of the "outperformance" of the optimal control with respect to the constant mix strategy and the portfolio fully invested in risky asset depends on the value of $\mu - r$ and σ as well. The benefits of the optimal control with respect to the constant versions can also be observed on Figure 3.2. The latter presents the empirical distributions of terminal wealth[10] for the strategies implementing the optimal control p^* and the constant control 0.5 when $\alpha = 0$. Interestingly, for the two setups with high risk premium and medium/low volatility as well as the configuration low risk premium low volatility, the optimal control implies a much fatter right tail. The combination of parameters leads to optimal control values more aggressive than the constant value of 0.5. This is no longer the case in the three remaining configurations. The distributions are much more concentrated around the initial wealth for the optimal control (lower allocation to the risky asset), leading to an important reduction of the downside risk.

The results of the strategies for the CRRA investor using p^* with $\alpha = 0.8$ are presented in Table 3.9. As expected, the values of terminal utility and Certainty Equivalent obtained with the optimal control are greater than those obtained with constant control equal to 5 and 10. Overall, the average optimal allocations to the risky asset are lower when the

[10] All the distributions are produced using a standard kernel density estimation.

capital protection is used, ranging from 97% to 11%, with respect to optimal allocations when $\alpha = 0$. Note that the ranges of optimal control (presented at the bottom of Table 3.9) are quite different depending on the parameters of the risky asset dynamics (above 9 for the most favorable setup versus less than 1 for the least one). The impact in terms of return can be observed on Figure 3.3, showing empirical distributions of terminal wealth for the strategies implementing the optimal control p^* and the constant control $\pi = 5$ when $\alpha = 0.8$. For optimal controls, the densities are more spiky than their equivalents without protection. Indeed, the drawdown constraint "cuts" the distribution below 0.8 and attenuates the weight of extreme gains. The outperformance of the optimal control with respect to the constant value is translated into a shift of the distribution toward wealth greater than one.

Interestingly, when we compare the results of the strategies using the optimal control with and without capital protection (Table 3.8 versus Table 3.9), a general conclusion on the preference of the CRRA investor cannot be drawn, even with optimal controls. In four out of six market configurations, the CE obtained with the capital protection in Table 3.9 is higher than or equal to the CE obtained with $\alpha = 0$ in Table 3.8. However for the two most favorable market parameter sets ($\mu - r = 0.07, \sigma = 0.1$ and $\mu - r = 0.07, \sigma = 0.2$), the gains in terms of average utility and Certainty Equivalent are larger without capital protection. Note that the only cases where the optimal strategy with capital protection leads to larger annualized returns and Sharpe ratios correspond to the least favorable market conditions with low risk premium. As expected the strategies with $\alpha = 0.8$ exhibit lower maximum drawdown in all setups (except the least favorable: average maximum drawdown of 8.23% versus 8.18%), and higher minimal values of annualized return. Again, in that sense, the contrast between Figures 3.2 and 3.3 is clear. The main differences between the two sets of distributions with optimal control concern clearly the tails of the distributions. Indeed, the left tail is flat below 0.8 in each configuration when the capital protection is implemented (Figure 3.3), which is not the case without drawdown constraint. As expected given the results discussed above, the right tail of the distributions with $\alpha = 0$ is fatter than the the case $\alpha = 0.8$ for the two most favorable market configurations (top left and middle left panels). Indeed, we can observe that larger values of wealth are obtained with $\mu - r = 0.07, \sigma = 0.1$ and $\mu - r = 0.07, \sigma = 0.2$, when the drawdown constraint is not present.

The following paragraphs are dedicated to the discussion of the results of the strategies implementing optimal control p^* for the prospect theory investor. In Table 3.10, we present the results with $\alpha = 0$ in terms of terminal utility. All Certainty Equivalent values are strictly greater than those obtained by the corresponding strategy with constant control equal to 0.5 and 1. The superiority of the optimal allocation program is also visible in terms of risk-adjusted performance: superior Sharpe ratio in all configurations with respect to $\pi = 0.5$. We can observe that the average optimal allocations to the risky asset are larger overall for this investor than for the CRRA in Table 3.8. Indeed, the average weight of stocks represents more than 100% in the four most favorable market configurations, whereas it is clearly reduced for the two remaining setups. This last result

is in line with Gomes (2005): loss-averse investors will not hold a large fraction of stocks unless the equity premium is quite high and the volatility quite low. When we compare the annualized returns obtained by the S-shape utility and the CRRA investors with their respective optimal control, we do not observe a superiority of one type of investor in all market configurations. In the two opposite extreme setups (high premium, low volatility and low premium, high volatility), the CRRA investor manages to obtain better absolute performance (even if the figures are rather close with the S-shape in the first case), whereas the results are more mitigated with the other risky asset parameter sets. It should be noted that the volatility of returns is much larger for the prospect theory investor, leading to more extreme observations in some scenarios. The empirical kurtosis measures confirm this observation.

Figure 3.4, presenting the terminal wealth distribution for the S-shape utility investor without drawdown constraint, allows to further analyze the profile of the resulting performance. First, the shape of densities corresponding to optimal control (continuous lines) are clearly different from those of the constant mix strategy ($\pi = 0.5$). In particular, for three configurations, the optimal control leads to a fatter right tail (high risk premium with low/medium volatility, and low risk premium with low volatility). For the remaining panels, the distribution exhibits a clear spike around one, i.e., the initial wealth and reference point Θ. The spikes are even more pronounced than for the corresponding CRRA densities (cf., Figure 3.2), showing the importance of Θ in the optimization integrating the S-shape utility.

Table 3.11 exhibits the results of optimal allocation programs with capital protection. Again, the benefits of using the optimal control instead of the constant values 5 or 10 are obvious in terms of mean terminal utility and Certainty Equivalent. Note that the (average) optimal allocation to the risky asset is never above 100% and are quite similar to the weights obtained for the CRRA investor (Table 3.9), except for the low risk premium configurations. In addition, we remark that the performance in terms of annualized returns are better for the investor with power utility. In fact, for all setups the latter investor obtains a greater (mean) annualized return (except the case high risk premium, low volatility: 7.57% versus 7.58%).

Furthermore, an important result can be drawn for the prospect theory investor when comparing Table 3.10 and 3.11. For all market configurations, using optimal controls, the strategies without protection lead to better results than strategies controlling drawdown. We can clearly observe that the Constant Equivalent and average terminal utility are always larger when $\alpha = 0$. This preference can be justified when we look at the difference of mean annualized returns: the strategy without protection dominates the portfolio insurance strategy.

As proposed earlier for the CRRA, we can also compare the resulting wealth densities for the prospect theory investor with and without protection. The distributions corresponding to $\alpha = 0.8$ are presented in Figure 3.5 and have to be compared with Figure 3.4, discussed above. We can clearly distinguish the presence of a mode at 1, i.e., the value of the reference point Θ, in the distributions of both types of strategies. Hence, as suggested

by Berkelaar et al. (2004), the reference level of the S-shape utility acts as an implicit protection, leading to a concentration of final wealth around the initial value of 1 even in the absence of any constraint. Therefore, the prospect theory investor invests wealth in order to insure the reference point Θ, with or without constraint, like a "partial portfolio insurance" as discussed by Berkelaar et al. (2004). The important difference is for levels of wealth below the reference point: in the absence of constraint, this investor is more aggressive (due to the convexity of the utility), implying potential large losses. This latter tail risk is limited when the drawdown constraint is implemented. As a consequence, in Figure 3.5 corresponding to the case $\alpha = 0.8$, we can even observe bi-modal densities with maximum at 0.8 and 1. Therefore the capital protection, in the prospect theory case, can be perceived as an additional constraint, which is not useful according to our results, given the presence of the reference point Θ. Note that the distributions obtained for the S-shape utility investor are clearly more spiky than their equivalents for the power utility investor, highlighting the loss-averse behavior of the first investor.

3.7.3 Prospect theory investor: impact of a higher reference point

In this section, we discuss the results obtained for a prospect theory investor with a higher reference point. Specifically, we set $\Theta = 1.05$ in the utility function given by equation (3.3). The remaining parameters have the same numerical values as in the previous section. This additional analysis allows us to investigate two aspects of the optimal behavior of the prospect theory investor. First, we can verify that the results obtained previously with $\Theta = 1$ in terms of shape of terminal wealth and strategy preference are still valid when the reference point is no longer equal to the initial level of wealth. Secondly, we can analyze the impact of setting a higher reference level in terms of performance. Note that the value 1.05 is reasonably attainable given the market parameters we use, nevertheless the investor has to take risk in order to ensure the discounted level of wealth Θ, as it cannot be reached with a full investment in the risk free asset.

In Table 3.12, we present the statistics related to the case $\alpha = 0$. A first observation concerns the average annualized rates of return with respect to the case $\Theta = 1$ (Table 3.10). Indeed, the figures are quite similar between the two investors, except for the case low risk premium high volatility. The investor with $\Theta = 1.05$ displays a lower average annualized return (-0.27% versus 0.03%) together with a larger (average) standard deviation (5.85 versus 1.97). Note also that the average maximum drawdown of the optimal strategy with $\Theta = 1.05$ are larger in the two least favorable market configurations: low risk premium high volatility (14.14% versus 4.36%) and low risk premium medium volatility (17.63% versus 13.87%).

The implications in terms of terminal wealth can be observed on Figure 3.6. The latter displays the distributions of final wealth values for the prospect theory investors with $\Theta = 1$ and $\Theta = 1.05$, implementing their respective optimal control. Due to the proximity between the two values of the reference level, the differences between the two types of curves are only slight. Interestingly, we can clearly note a similar pattern between the two optimal strategies. In fact, the shape of the distributions are similar, but the

curves corresponding to the case $\Theta = 1.05$ seem to be shifted toward more positive values. Indeed, the reference point acts, as described in the previous section, as an hypothetical level of protection whether $\Theta = 1$ or $\Theta = 1.05$. Hence the spikes in the distributions imply a larger mode around 1.05 for the latter case rather than 1. The optimal behavior of the prospect theory investor, consisting in trying to ensure the value of wealth Θ, is confirmed. In addition, we observe that the fat left tails induced by the risk-seeking behavior below Θ are present for both investors.

The results for the investor with $\Theta = 1.05$ implementing the optimal control in the presence of the drawdown constraint are presented in Table 3.13. The comparison in terms of annualized return with the investor $\Theta = 1$ (Table 3.11) leads to the same conclusions, whether or not the protection is implemented. Indeed, the rate of returns are again similar between the two investors implementing the optimal control when $\alpha = 0.8$, except for the two cases low premium high volatility and low premium medium volatility, where the investor with $\Theta = 1.05$ obtains lower performance. Again, for these two setups, the latter investor also displays a notably larger volatility of annualized return as well as a larger magnitude of maximum drawdown. The reason for that is the higher level of Θ: given that the reference point is greater the investor will take more risk (i.e., larger "zone" where the risk-seeking behavior prevails) in order to reach this level of wealth. The implications can be seen in terms of average allocation to the risky asset in the two market configurations mentioned above: respectively 14% and 9% when $\Theta = 1.05$ versus 7% and 1% when $\Theta = 1$.

As for the case $\alpha = 0$, we present in Figure 3.7, the terminal wealth for the investors with $\Theta = 1$ and $\Theta = 1.05$ implementing the optimal control when the drawdown constraint is implemented ($\alpha = 0.8$). The conclusions on the distributions for the case $\alpha = 0$ remain valid. Indeed, the curves are very similar between the two investors, while the distributions are shifted toward the right when $\Theta = 1.05$. We observe clearly bimodal densities in most setups, with modes around 0.8 and at the value of the reference point. Interestingly, for the cases low premium medium volatility (middle right panel) and low premium high volatility (bottom right panel), the densities corresponding to the investor with $\Theta = 1.05$ exhibit a larger fraction of wealth just above the boundary 0.8 (implied by the drawdown constraint). Indeed, for these two configurations, the investor with $\Theta = 1.05$ is more aggressive in order to reach the reference point even if market conditions are not favorable. This additional risk-taking naturally leads to more "crash scenarios" when the drawdown constraint limits the losses.

We focus now on the comparison between the optimal strategies with $\alpha = 0$ and $\alpha = 0.8$ when the reference point Θ is set to 1.05 in both cases. Clearly comparing Tables 3.12 and 3.13, we can observe that the strategy without drawdown constraint leads to larger values of Constant Equivalent. Indeed, in all market configurations, the average terminal utility is strictly greater when the drawdown constraint is absent. Of course, these results are achieved with larger allocations to the risky assets.

Finally, whether or not the reference point of the prospect theory investor is higher than the initial level of wealth, the conclusion remains the same. The optimal allocation

program of this type of investor is: trying to reach the amount Θ by taking risk (risk-seeking behavior) and then, if this objective is achieved, trying to maintain the level of wealth above Θ through a defensive allocation when the level of wealth gets closer to the reference point. Due to the fact that this behavior persists whether or not the drawdown limitation is implemented, the latter constraint does not appear to lead to gains in terms of utility when the optimal strategy is implemented.

3.7.4 Mis-estimation risk: impact on preference

The parameters of the risky asset generating process are key inputs in our optimization procedure as illustrated by the expression of the control p^* in equation (3.20). In the previous section, these parameters are assumed to be known throughout our analysis in order to derive optimal control values. However in practice, the estimation of the "true" volatility and expected returns of stocks can be a difficult task[11]. In order to evaluate the impact of mis-estimates of these parameters, we discuss below the results of strategies with and without capital protection, implementing "optimal" controls p^* derived using a mis-estimate of the risk premium or volatility. The results show that the drawdown constraint can be perceived as an insurance against model's mis-estimation risk.

We focus first on the results for the CRRA investor with power utility. In Tables 3.14 and 3.15, we present the results of strategies respectively with $\alpha = 0$ and $\alpha = 0.8$, using an optimal control derived with underestimation or overestimation of the risk premium $\hat{\mu} - r$. The true value of the volatility is assumed to be known. Hence, the portfolio sample paths are generated using the control p^* derived with an estimate of the risk premium $\hat{\mu} - r = 0.07$ (resp. 0.02) while the risky asset trajectories are generated with the parameter $\mu - r = 0.02$ (resp. 0.07). Interestingly, in five configurations out of six, the strategies with capital protection lead to larger values of terminal utility. The only case where the investment without protection is preferred corresponds to an underestimation of the risk premium (0.02 instead of 0.07 for the true value) together with a volatility equal to 0.1. The usefulness of the capital protection is clearly highlighted in setups with overestimation of the risk premium (right part of Tables 3.14 and 3.15). The average maximum drawdown of the strategies without protection lays between 26% and 39% for these scenarios (versus a range between 16% and 17% with protection). The benefits can also be observed in terms of standard deviation of annualized returns. Remark that, in the case without constraint, the strategies with optimal control with mis-estimate of the risk premium do not outperform the constant control in all setups. In fact, except in the two opposite extreme scenarios, the constant control 0.5 leads to larger CE than the optimal one when the true value of $\mu - r$ is unknown. On the other hand, when the drawdown constraint is considered, even if the risk premium is mis-estimated, the "optimal" portfolio achieves a larger utility in most setups with respect to constant control values of 5 and 10.

Figures 3.8 and 3.9 illustrate the impact of mis-estimation of the risk premium in terms of terminal wealth distribution, respectively without and with drawdown constraint. Each

[11] See for instance Merton (1980) on the estimation of expected returns.

panel shows the distribution of the optimal strategy using the true values of parameters (continuous line) as well as the strategy optimized with an overestimate or underestimate of $\mu - r$. Clearly in case of overestimation of the risk premium (right panels), the "crash risk" is important when the protection is not considered, as illustrated by the fat left tails of strategies using mis-estimates in Figure 3.8 (whereas the left tails remain cut under 0.8 with the constraint in Figure 3.9). On the other hand, the opportunity cost implied by the underestimation of $\mu - r$ is present both with and without protection (cf., right tails of distributions).

We run the same type of exercise with over and underestimation of the volatility. More specifically, the controls p^* are derived with an estimate $\hat{\sigma}$, different from the true volatility parameter σ, while the value of $\mu - r$ is assumed to be known. The results are presented in Tables 3.16 and 3.17, respectively for the protections $\alpha = 0$ and $\alpha = 0.8$. For each risk premium, we consider two "true" values of σ equal to 0.1 and 0.3, respectively with overestimation ($\hat{\sigma} = 0.2$ and $\hat{\sigma} = 0.3$), and underestimation ($\hat{\sigma} = 0.2$ and $\hat{\sigma} = 0.3$), leading to height different setups. As in the case of mis-estimation of the risk premium, in most configurations (six out of eight), the strategy with $\alpha = 0.8$ leads to a greater average terminal utility than the strategy without protection. We remark that the setups implying a larger utility for the strategy with $\alpha = 0$ correspond to an overestimation of the risky asset volatility when $\mu - r = 0.07$. In addition, it can be noted that the strategy without protection leads to a remarkably low Constant Equivalent of 0.33 when the risk premium is high and the volatility is largely underestimated ($\hat{\sigma} = 0.3$ versus $\sigma = 0.1$).

Therefore, even if this exercise with parameters mis-estimate does not dramatically change the resuls obtained previously for the CRRA investor, it should be noted that the capital protection $\alpha = 0.8$ can be useful in specific market conditions with overestimation of the risk premium or underestimation of volatility. Again, in the most favorable configurations, the investor with power utility appears to prefer the strategy with $\alpha = 0$, emphasizing the opportunity cost implied by the drawdown constraint.

Similarly, we discuss now the results obtained for the investor with S-shape utility using under or overestimates of the risky asset parameters in the optimization procedure. The results corresponding to mis-estimation of the risk premium are presented in Tables 3.18 and 3.19, respectively without and with capital protection. Unlike for the CRRA, the prospect theory investor prefer the strategy with $\alpha = 0$ in most configurations. Indeed, except when the risk premium is overestimated with volatility equal to 0.20 or 0.30, the Constant Equivalent values are larger in Table 3.18 corresponding to the absence of drawdown constraint. Note that for the two latter configurations, the strategy with $\alpha = 0$ exhibits very severe maximum drawdown: respectively 65% and 70% on average when the volatility is equal to 0.20 and 0.30. However, unlike for the CRRA, in terms of terminal utility, the prospect theory investor strongly prefers the strategy without protection in case of underestimation of the risk premium for all values of volatility.

We further illustrate the impact of mis-estimation of the risk premium in terms of terminal wealth distribution in Figures 3.10 and 3.11, respectively without and with drawdown constraint. As for the CRRA above, each panel shows the distribution of

the optimal strategy using the true values of parameters (continuous line) as well as the optimized strategy using an overestimate or underestimate of $\mu - r$ (dashed line). In case of overestimation of the risk premium (right panels), it appears that the "crash risk" seems to be still important for the prospect theory investor when the protection is not used (cf., left tails in Figure 3.10). Nevertheless, the opportunity cost appearing when $\mu - r$ is underestimated, is much less visible than for the CRRA investor, as shown by the top left panel of Figure 3.10. Note that the optimal portfolios of the prospect theory investor still exhibit the importance of the reference point Θ and the implicit objective of insuring this amount of wealth.

When we run the same exercise with over and underestimation of the volatility, the results are even more in favor of the strategy without capital protection. Tables 3.20 and 3.21 clearly illustrate this observation. In Table 3.20 corresponding to $\alpha = 0$, the Constant Equivalent measures are always larger than the corresponding values in Table 3.21 with $\alpha = 0.8$, except in one setup (corresponding to a large underestimation of volatility: $\hat{\sigma} = 0.3$ versus $\sigma = 0.1$). This result contrasts with the preference of the CRRA investor described above. As expected the average allocation to the risky asset are always lower in the presence of the capital protection. Therefore the tendency to prefer optimal strategies without drawdown constraint remains valid overall for the prospect theory investor even in the presence of mis-estimation risk. This further puts forward the difficulty to bring forward a potential usefulness of capital protection in a prospect theory context where optimal portfolios are implemented, even in the presence of mis-estimation of risk. The only configurations where the drawdown constraint could be gainful for this investor are: overestimation of the risk premium when the volatility is at a substantial level, and severe underestimation of the volatility when the risk premium is low. In these specific cases, the drawdown constraint can act as an insurance against model's mis-estimation risk.

3.7.5 Shorter investment horizon: impact on preference

Changing the investment horizon can affect the way an investor evaluates an investment strategy[12]. In particular, long horizons can change the optimal portfolio allocation as discussed by Barberis (2000) and Pástor and Stambaugh (2012), especially for loss averse investors (cf., Berkelaar et al., 2004). In order to figure out if our previous results are substantially modified with a shorter investment horizon, we present below a similar analysis with an investment horizon T of one year[13]. As with the 5-year horizon discussed above, for each investor, portfolios implementing optimal controls lead to greater average terminal utility and Constant Equivalent than their equivalent with constant control (unreported for space consideration). In this section, we focus our analysis on the comparison between 1-year optimal portfolios (with and without constraint) and 5-year optimal strategies.

Tables 3.22 and 3.23 present the results of optimal portfolios for the CRRA investor,

[12]Under specific assumptions, the portfolio selection problem does not depend on the investment horizon, cf., Samuelson (1963), Ross (1999) and De Brouwer and Van den Spiegel (2001) on that issue.

[13]Using constant control, Dierkes et al. (2010) find that portfolio insurance is attractive for prospect theory investors for almost every investment horizon between one month and seven years.

respectively with $\alpha = 0$ and $\alpha = 0.8$, with an investment horizon of one year. Focusing on Constant Equivalent values, we can see that in most configurations, the optimal portfolio with drawdown constraint leads to a larger CE than the portfolio with $\alpha = 0$. In fact, the only setup where the last statement is not verified is the most favorable with high risk premium and low volatility. These observations are also observed in terms of annualized return. Note that the average Calmar ratios of the protected strategies are all greater than their equivalent when $\alpha = 0$.

Without drawdown constraint, whereas the horizon does not have an important impact in terms of allocation to the risky asset due to the myopic behavior of the CRRA investor (cf., average stock allocation in Table 3.8 when $T = 5$), we remark that the Constant Equivalents are larger when the horizon is equal to five years (as expected given the positive drift of the risky asset dynamics) while annualized returns are higher for the shortest horizon. Note also that the average values of maximum drawdown are clearly higher in magnitude for the longest horizon. This is highlighted in Figure 3.12, comparing the terminal wealth distributions of optimal portfolios with horizon of five and one year (with $\alpha = 0$). As expected the tails of the 5-year distributions are fatter than for the one-year portfolios. However, the density just below the initial level of wealth (equal to one), seems to display a heavier weight for the shortest horizon. In particular, this is striking in the upper left panel. Note that the range of annualized returns (and their standard deviation) is always larger for the short horizon (even if it is not necessarily clearly visible on Figure 3.12 presenting terminal wealth). This highlights the important risk of extreme "bullish or bearish crash" when $T = 1$.

The previous comment regarding allocation to the risky asset is no longer valid when the drawdown constraint is implemented. Indeed, we observe that the optimal portfolios with short horizon have a greater average allocation to stocks in all setups with respect to strategies with 5-year horizon in Table 3.9. Interestingly, these larger risky allocations do not imply worse maximum drawdown (thereby the Calmar ratios are better when $T = 1$). In that sense, the distributions presented in Figure 3.13 are very informative. As discussed above, they correspond to terminal wealth under drawdown constraint with horizon of five and one year. The divergence between the two types of distributions are much less visible with respect to the case $\alpha = 0$ in Figure 3.12. Of course, despite this proximity more pronounced, the shape of distribution remains more spiky for the 1-year portfolios.

We discuss now the equivalent results for the investor with S-shape utility. First, we compare Table 3.24 and 3.25, presenting metrics related to 1-year optimal portfolios respectively without and with drawdown constraint. As for the 5-year horizon discussed in section 3.7.2, we observe that the Constant Equivalents values are all larger for the portfolio without constraint with respect to $\alpha = 0.8$ (note though that four setups lead to the same level of average terminal utility). Even in configurations with low risk premium, the protection penalizes the portfolio in terms of annualized returns, even if it leads to lower volatility.

Unlike for the CRRA, reducing the investment horizon from five years to one year implies a clear reduction in the optimal allocation to the risky asset when $\alpha = 0$ for

the prospect theory investor, as suggested by Berkelaar et al. (2004). Indeed, when we compare Tables 3.10 (5-year) and 3.24 (1-year), we remark that the average allocation to stocks is always lower in the latter. Remark that this lower allocation implies a lower average returns only when the risk premium is low. An additional implication of this lower allocation is that in four configurations (out of six), the portfolios with short horizon display a lower standard deviation of annualized returns.

The impact of the horizon on the terminal wealth distributions are even more impressive in Figure 3.14, comparing optimal portfolios with $\alpha = 0$ with horizon equal to 5 and 1 year, for the prospect theory investor. The importance of the reference point Θ (equal to one) implies very spiky densities for the short horizon. In particular, this effect is very obvious when the risk premium is low.

Finally, we discuss the divergences between 5-year and 1-year optimal portfolios when the drawdown constraint is implemented with the S-shape utility (Table 3.11: 5-year, and Table 3.25: 1-year). Similarly to the case $\alpha = 0$, reducing the investment horizon implies a reduction in the optimal allocation in most market configurations (except when the risk premium is high with low/medium volatility). Hence, except in the two latter cases, the 5-year horizon portfolios lead to greater (average) annualized rate of return. Remark that the magnitude of maximum drawdown are always larger in Table 3.11 when the horizon is the longest.

The conclusions regarding the distributions of terminal wealth for optimal portfolios with protection are similar to the case without constraint. In Figure 3.15, we compare optimal portfolios with $\alpha = 0.8$ with horizon equal to 5 and 1 year, for the prospect theory investor. As expected, the densities corresponding to the shortest horizon are very spiky, and exhibit a clear mode around one, emphasizing again the role of Θ. This confirms that, even when the horizon is quite short, the optimal portfolio for the S-shape utility, is managed in order to ensure the reference point. Hence, the presence of the non-zero α appears to be more a constraint for the upside, rather than a benefit for the downside risk protection.

Conclusion

In this paper we analyze optimal investment strategies under the drawdown constraint that the wealth process never falls below a fixed fraction of its running maximum. In particular, from an expected utility standpoint, we address the question of the preference for optimal strategies implementing the drawdown constraint or not. We study the finite horizon expectation maximization problem for investors with power utility (CRRA) as well as S-shape utility function, asymmetric over gains and losses relative to the reference point (prospect theory). The optimal portfolio allocations are calculated by solving numerically the Hamilton-Jacobi-Bellman equation associated with the dynamic programming principle related to the stochastic control problem.

Our main results are based on stochastic simulations of optimal portfolios with and without drawdown constraints. More specifically, making vary the market model param-

eters, we simulate risky asset scenarios as well as optimal portfolios, and we compare ex-post results in terms of average utility, returns and risks. Our simulations are performed with investment horizon of five years. We also consider a shorter horizon of one year, and a case where parameters of the risky asset generating process are mis-estimated.

For the CRRA investor, we find that implementing the optimal portfolio with drawdown constraint can be gainful in terms of utility in certain specific market conditions such as low risk premium and/or high volatility for both the 5-year and 1-year horizons. In addition, when the risk premium is overestimated or the volatility underestimated, the "optimized" portfolio with mis-estimates implementing the drawdown constraint can lead to a larger utility for this investor.

The last statement is also true for the S-shape utility. However, for this type of investor, when the parameters of the risky asset data generating process are assumed to be known, the optimal strategy without constraint systematically leads to better results in terms of utility than the equivalent optimal strategy with drawdown control. This result holds with both the 5-year and 1-year investment horizon. A main reason is that in an optimal context, with or without drawdown constraint, the prospect theory investor manages the portfolio in order to insure the reference point Θ. The latter parameter then acts as an "implicit" protection level. Therefore, a main conclusion of this paper, is that, using optimal allocation, the potential usefulness of strategies controlling drawdown appears to be clearly reduced in a prospect theory context.

Given that in practice, management fees for implementing strategies with drawdown constraint are nonzero, a natural extension of this paper would be to study utility indifference prices (see Barrieu and El Karoui, 2009) for different kind of investors. Indeed, the derivation of the amount the investor is willing to accept in order to buy/sell the strategy, and its comparison with actual fees, could lead to a better understanding of the "true cost" of implementing the drawdown constraint in portfolios for different investors.

Tables and Figures

	Risk premium	
Volatility	Low	High
Low	$\mu - r = 0.02$ $\sigma = 0.1$	$\mu - r = 0.07$ $\sigma = 0.1$
Medium	$\mu - r = 0.02$ $\sigma = 0.2$	$\mu - r = 0.07$ $\sigma = 0.2$
High	$\mu - r = 0.02$ $\sigma = 0.3$	$\mu - r = 0.07$ $\sigma = 0.3$

Table 3.1: Summary of the 6 stock market configurations

Stat., $\alpha = 0, \pi = 0.5$	$\mu - r = 0.07$			$\mu - r = 0.02$		
	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$
Mean Ann. return (%)	3.46	3.16	2.65	0.91	0.61	0.11
Min Ann. return (%)	-5.5	-13.98	-21.9	-7.83	-16.1	-23.82
Max Ann. return (%)	13.97	25.1	36.99	11.15	22.02	33.6
Std. dev. Ann. return	2.32	4.63	6.91	2.26	4.52	6.74
Mean Ann. volatility (%)	5	10	15	5	10	15
Min Ann. volatility (%)	4.57	9.14	13.71	4.57	9.14	13.71
Max Ann. volatility (%)	5.41	10.82	16.23	5.41	10.82	16.22
Skewness	0.01	0.02	0.03	0.01	0.02	0.03
Kurtosis	0	0	0	0	0	0
Max. drawdown (%)	8.31	19.44	29.78	11.2	22.69	32.86
Sharpe ratio	0.09	0.02	-0.02	-0.42	-0.24	-0.19
Calmar ratio	0.53	0.25	0.16	0.16	0.09	0.07
Average stock allocation (%)	50	50	50	50	50	50
Stat., $\alpha = 0, \pi = 1$	$\mu - r = 0.07$			$\mu - r = 0.02$		
	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$
Mean Ann. return (%)	6.83	5.57	3.49	1.62	0.42	-1.56
Min Ann. return (%)	-10.92	-26.74	-40.35	-15.26	-30.31	-43.26
Max Ann. return (%)	29.56	54.95	83.48	23.24	47.4	74.53
Std. dev. Ann. return	4.8	9.49	13.98	4.56	9.03	13.3
Mean Ann. volatility (%)	10	20	30.01	10	20	30
Min Ann. volatility (%)	9.14	18.28	27.42	9.14	18.28	27.42
Max Ann. volatility (%)	10.82	21.63	32.44	10.82	21.63	32.44
Skewness	0.02	0.04	0.06	0.02	0.04	0.06
Kurtosis	0	0	0	0	0	0
Max. drawdown (%)	16.03	35.54	51.5	21.29	40.68	55.69
Sharpe ratio	0.68	0.28	0.12	0.16	0.02	-0.05
Calmar ratio	0.54	0.24	0.14	0.15	0.07	0.03
Average stock allocation (%)	100	100	100	100	100	100

Table 3.2: Statistics over 50,000 draws of portfolios with constant control. The protection level α is set to 0. The control is set to 0.5 in the upper panel, implying that the strategy is a pure constant mix, i.e., a combination of 50% of risky asset and 50% of riskless security, while it is set to 1 the bottom panel, corresponding to a pure investment in the risky asset. All amounts are simulated over 5 years and evaluated in terms of their discounted values.

Stat., $\alpha = 0.8, \pi = 5$	$\mu - r = 0.07$			$\mu - r = 0.02$		
	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$
Mean Ann. return (%)	5.37	3.01	1.69	1.31	0.79	0.53
Min Ann. return (%)	-3.89	-4.36	-4.36	-4.23	-4.36	-4.36
Max Ann. return (%)	27.27	44.1	55.53	21.09	37.1	47.99
Std. dev. Ann. return	4.22	5.55	5.43	3.25	4.23	4.44
Mean Ann. volatility (%)	7.27	9.13	9.17	6.25	7.98	8.36
Min Ann. volatility (%)	2.76	2.16	2.38	2.04	2.11	2.36
Max Ann. volatility (%)	9.16	16.19	21.52	8.95	15.78	20.9
Skewness	0.16	0.35	0.66	0.13	0.32	0.62
Kurtosis	0.78	6.38	19.74	1.61	9.08	23.43
Max. drawdown (%)	11.06	16.18	16.98	13.15	16.73	17.09
Sharpe ratio	0.68	0.2	0.03	0.13	-0.03	-0.08
Calmar ratio	0.57	0.21	0.11	0.14	0.06	0.04
Average stock allocation (%)	70.19	38.42	20.32	59.09	31.62	17.29
Stat., $\alpha = 0.8, \pi = 10$	$\mu - r = 0.07$			$\mu - r = 0.02$		
	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$
Mean Ann. return (%)	7.58	2.27	1	1.57	0.64	0.41
Min Ann. return (%)	-4.35	-4.36	-4.36	-4.36	-4.36	-4.36
Max Ann. return (%)	54.45	71.56	58.39	39.86	55.32	48.26
Std. dev. Ann. return	7.46	6.23	5.13	4.74	4.68	4.51
Mean Ann. volatility (%)	10.85	9.8	9.03	8.42	8.61	8.53
Min Ann. volatility (%)	2.6	2.44	2.56	2.14	2.43	2.69
Max Ann. volatility (%)	16.64	26.1	28.97	15.94	24.71	27.28
Skewness	0.37	1.07	1.83	0.34	0.99	1.71
Kurtosis	3.68	34.36	88.59	7.93	42.21	95.49
Max. drawdown (%)	15.29	16.78	16.3	16.53	16.85	16.34
Sharpe ratio	0.58	0.06	-0.05	0.05	-0.07	-0.1
Calmar ratio	0.54	0.14	0.07	0.11	0.04	0.03
Average stock allocation (%)	96.49	27.13	10.16	68.54	21.18	9

Table 3.3: Statistics over 50,000 draws of portfolios with constant control. The protection level α is set to 0.8. The control is set to 5 in the upper panel and to 10 in the lower panel. All amounts are simulated over 5 years and evaluated in terms of their discounted values.

CRRA, $\alpha = 0, \pi = 0.5$	$\mu - r = 0.07$			$\mu - r = 0.02$		
	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$
Constant Equivalent	1.18	1.13	1.06	1.04	1	0.94
Mean terminal utility	-0.85	-0.88	-0.94	-0.96	-1	-1.06
Std. dev. terminal utility	0.1	0.2	0.33	0.11	0.23	0.37
Qt. 1% terminal utility	-1.1	-1.45	-1.95	-1.24	-1.65	-2.21
Qt. 50% terminal utility	-0.84	-0.86	-0.89	-0.96	-0.97	-1
Qt. 99% terminal utility	-0.65	-0.51	-0.41	-0.74	-0.58	-0.46
CRRA, $\alpha = 0, \pi = 1$	$\mu - r = 0.07$			$\mu - r = 0.02$		
	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$
Constant Equivalent	1.35	1.16	0.9	1.05	0.9	0.7
Mean terminal utility	-0.74	-0.86	-1.11	-0.95	-1.11	-1.42
Std. dev. terminal utility	0.17	0.41	0.84	0.22	0.52	1.08
Qt. 1% terminal utility	-1.22	-2.22	-4.25	-1.57	-2.85	-5.46
Qt. 50% terminal utility	-0.72	-0.78	-0.88	-0.93	-1	-1.13
Qt. 99% terminal utility	-0.43	-0.28	-0.19	-0.55	-0.35	-0.24

Table 3.4: Performance measurements for CRRA (2) investor over 50,000 draws of portfolios. The protection level α is set to 0. The control is set to 0.5 in the upper panel, implying that the strategy is a pure constant mix, i.e., a combination of 50% of risky asset and 50% of riskless security, while it is set to 1 the bottom panel, corresponding to a pure investment in the risky asset. All amounts are simulated over 5 years and evaluated in terms of their discounted values.

CRRA, $\alpha = 0.8, \pi = 5$	$\mu - r = 0.07$			$\mu - r = 0.02$		
	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$
Constant Equivalent	1.27	1.12	1.05	1.05	1.02	1
Mean terminal utility	-0.79	-0.9	-0.95	-0.95	-0.98	-1
Std. dev. terminal utility	0.15	0.21	0.21	0.14	0.18	0.18
Qt. 1% terminal utility	-1.13	-1.22	-1.25	-1.19	-1.24	-1.25
Qt. 50% terminal utility	-0.78	-0.92	-1	-0.96	-1.02	-1.04
Qt. 99% terminal utility	-0.47	-0.39	-0.39	-0.6	-0.49	-0.46
CRRA, $\alpha = 0.8, \pi = 10$	$\mu - r = 0.07$			$\mu - r = 0.02$		
	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$
Constant Equivalent	1.35	1.07	1.02	1.05	1.01	1
Mean terminal utility	-0.74	-0.93	-0.98	-0.95	-1	-1
Std. dev. terminal utility	0.23	0.22	0.2	0.19	0.19	0.18
Qt. 1% terminal utility	-1.19	-1.25	-1.25	-1.23	-1.25	-1.25
Qt. 50% terminal utility	-0.73	-0.98	-1.02	-0.99	-1.04	-1.04
Qt. 99% terminal utility	-0.28	-0.33	-0.4	-0.45	-0.44	-0.45

Table 3.5: Performance measurements for CRRA (2) investor over 50,000 draws of portfolios. The protection level α is set to 0.8. The control is constant and set to 5 in the upper panel and 10 in the lower panel. All amounts are simulated over 5 years and evaluated in terms of their discounted values.

S-shape, $\alpha = 0, \pi = 0.5$	$\mu - r = 0.07$			$\mu - r = 0.02$		
	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$
Constant Equivalent	1.18	1.13	1.07	1.01	0.99	0.97
Mean terminal utility	0.22	0.16	0.1	0.02	-0.05	-0.11
Std. dev. terminal utility	0.16	0.37	0.57	0.21	0.4	0.57
Qt. 1% terminal utility	-0.27	-0.81	-1.19	-0.54	-0.99	-1.32
Qt. 50% terminal utility	0.23	0.2	0.16	0.07	0.04	-0.02
Qt. 99% terminal utility	0.58	0.96	1.39	0.4	0.75	1.14
S-shape, $\alpha = 0, \pi = 1$	$\mu - r = 0.07$			$\mu - r = 0.02$		
	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$
Constant Equivalent	1.39	1.27	1.14	1.03	0.97	0.92
Mean terminal utility	0.44	0.32	0.17	0.04	-0.1	-0.23
Std. dev. terminal utility	0.34	0.77	1.17	0.39	0.74	1.06
Qt. 1% terminal utility	-0.5	-1.33	-1.78	-0.92	-1.54	-1.88
Qt. 50% terminal utility	0.43	0.33	0.17	0.11	0	-0.33
Qt. 99% terminal utility	1.28	2.34	3.66	0.83	1.7	2.77

Table 3.6: Performance measurements for S-shape utility investor over 50,000 draws of portfolios. The protection level α is set to 0. The control is set to 0.5 in the upper panel, implying that the strategy is a pure constant mix, i.e., a combination of 50% of risky asset and 50% of riskless security, while it is set to 1 the bottom panel, corresponding to a pure investment in the risky asset. All amounts are simulated over 5 years and evaluated in terms of their discounted values.

S-shape, $\alpha = 0.8, \pi = 5$	$\mu - r = 0.07$			$\mu - r = 0.02$		
	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$
Constant Equivalent	1.29	1.11	1.02	1.02	0.99	0.98
Mean terminal utility	0.34	0.15	0.04	0.03	-0.03	-0.06
Std. dev. terminal utility	0.3	0.43	0.44	0.27	0.35	0.37
Qt. 1% terminal utility	-0.34	-0.5	-0.54	-0.46	-0.53	-0.55
Qt. 50% terminal utility	0.33	0.11	0.01	0.05	-0.08	-0.13
Qt. 99% terminal utility	1.1	1.46	1.5	0.69	1.02	1.13
S-shape, $\alpha = 0.8, \pi = 10$	$\mu - r = 0.07$			$\mu - r = 0.02$		
	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$
Constant Equivalent	1.45	1.06	1	1.02	0.99	0.98
Mean terminal utility	0.5	0.09	-0.02	0.03	-0.05	-0.07
Std. dev. terminal utility	0.57	0.51	0.43	0.38	0.39	0.38
Qt. 1% terminal utility	-0.45	-0.55	-0.55	-0.52	-0.55	-0.55
Qt. 50% terminal utility	0.41	0.03	-0.08	0.02	-0.12	-0.14
Qt. 99% terminal utility	2.28	1.88	1.41	1.19	1.22	1.17

Table 3.7: Performance measurements for S-shape utility investor over 50,000 draws of portfolios. The protection level α is set to 0.8. The control is constant and set to 5 in the upper panel and 10 in the lower panel. All amounts are simulated over 5 years and evaluated in terms of their discounted values.

CRRA, $\alpha = 0, \pi = p^*$	$\mu - r = 0.07$			$\mu - r = 0.02$		
	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$
Constant Equivalent	1.65	1.17	1.07	1.05	1.01	1.01
Mean terminal utility	-0.61	-0.86	-0.93	-0.95	-0.99	-0.99
Std. dev. terminal utility	0.29	0.34	0.24	0.21	0.11	0.07
Qt. 1% terminal utility	-1.57	-1.95	-1.64	-1.55	-1.27	-1.18
Qt. 50% terminal utility	-0.55	-0.8	-0.9	-0.93	-0.98	-0.99
Qt. 99% terminal utility	-0.19	-0.33	-0.5	-0.56	-0.76	-0.84
Mean Ann. return (%)	13.22	4.96	2.18	1.6	0.4	0.18
Min Ann. return (%)	-21.43	-22.72	-16.79	-14.81	-8.11	-5.56
Max Ann. return (%)	51.14	47.63	27.61	22.93	10.42	6.73
Std. dev. Ann. return	10.17	8.09	5.25	4.47	2.21	1.47
Mean Ann. volatility (%)	20.01	17.13	11.43	9.8	4.9	3.27
Min Ann. volatility (%)	16.29	15.76	10.5	8.99	4.49	2.99
Max Ann. volatility (%)	21.64	18.54	12.36	10.59	5.29	3.53
Skewness	0.04	0.03	0.02	0.02	0.01	0.01
Kurtosis	0	0	0	0	0	0
Max. drawdown (%)	29.88	31.24	23.5	20.91	11.76	8.18
Sharpe ratio	0.66	0.29	0.19	0.16	0.08	0.05
Calmar ratio	0.55	0.24	0.17	0.15	0.1	0.09
Average stock allocation (%)	199.98	85.64	38.09	97.96	24.5	10.89
Qt. 1% average control	2	0.84	0.38	0.97	0.24	0.11
Qt. 50% average control	2	0.86	0.38	0.98	0.25	0.11
Qt. 99% average control	2	0.86	0.38	0.98	0.25	0.11

Table 3.8: Performance measurements for CRRA (2) investor over 50,000 draws of portfolios. The protection level α is set to 0. The optimal control is calculated numerically. All amounts are simulated over 5 years and evaluated in terms of their discounted values. The corresponding terminal wealth distributions are presented in Figure 3.2.

CRRA, $\alpha = 0.8, \pi = p^*$	$\mu - r = 0.07$			$\mu - r = 0.02$		
	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$
Constant Equivalent	1.36	1.13	1.07	1.06	1.02	1.01
Mean terminal utility	-0.74	-0.88	-0.93	-0.95	-0.98	-0.99
Std. dev. terminal utility	0.22	0.2	0.18	0.16	0.11	0.08
Qt. 1% terminal utility	-1.19	-1.22	-1.23	-1.22	-1.22	-1.18
Qt. 50% terminal utility	-0.73	-0.9	-0.94	-0.96	-0.98	-0.99
Qt. 99% terminal utility	-0.3	-0.43	-0.53	-0.58	-0.75	-0.83
Mean Ann. return (%)	7.57	3.22	1.87	1.48	0.52	0.25
Min Ann. return (%)	-4.25	-4.26	-4.26	-4.25	-4.26	-4.25
Max Ann. return (%)	36.51	35.72	26.21	22.43	11.23	7.19
Std. dev. Ann. return	7.08	5.27	4.13	3.66	2.23	1.52
Mean Ann. volatility (%)	10.58	9.12	7.83	7.1	4.75	3.3
Min Ann. volatility (%)	2.46	2.05	2	1.95	2.11	2.18
Max Ann. volatility (%)	15.23	13.49	10.46	9.28	5.28	3.62
Skewness	0.3	0.22	0.16	0.13	0.08	0.06
Kurtosis	2.28	1.97	1.17	0.94	0.13	0.05
Max. drawdown (%)	15.11	16.4	15.8	14.81	11.31	8.23
Sharpe ratio	0.61	0.25	0.15	0.13	0.09	0.07
Calmar ratio	0.55	0.23	0.16	0.15	0.11	0.1
Average stock allocation (%)	96.75	41.99	24.73	68.03	23.53	10.96
Qt. 1% average control	9	3.57	2	5.27	1.4	0.61
Qt. 50% average control	9.41	4.94	2.68	6.45	1.66	0.68
Qt. 99% average control	9.81	7.83	6.38	8.79	4.59	1.47

Table 3.9: Performance measurements for CRRA (2) investor over 50,000 draws of portfolios. The protection level α is set to 0.8. The optimal control is calculated numerically. All amounts are simulated over 5 years and evaluated in terms of their discounted values. The corresponding terminal wealth distributions are presented in Figure 3.3.

S-shape, $\alpha = 0, \pi = p^*$	$\mu - r = 0.07$			$\mu - r = 0.02$		
	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$
Constant Equivalent	1.94	1.6	1.24	1.07	1	1
Mean terminal utility	0.94	0.64	0.29	0.09	0.01	0
Std. dev. terminal utility	0.86	1.69	1.75	0.65	0.29	0.14
Qt. 1% terminal utility	-0.92	-1.92	-2.16	-1.45	-1.51	-0.04
Qt. 50% terminal utility	0.85	0.17	0.03	0.03	0.02	0.01
Qt. 99% terminal utility	3.5	5.37	5.38	1.89	0.62	0.11
Mean Ann. return (%)	13.21	7.72	1.97	2.13	0.21	0.03
Min Ann. return (%)	-21.43	-48.43	-60.79	-28.91	-42.46	-49.61
Max Ann. return (%)	51.35	52.17	52.36	50.27	50.4	12.95
Std. dev. Ann. return	10.18	18.81	20.76	8.09	3.78	1.97
Mean Ann. volatility (%)	19.99	38.97	44.81	17.24	6.22	1.92
Min Ann. volatility (%)	16.1	20.42	21.38	11.79	2.35	0.75
Max Ann. volatility (%)	21.62	43.26	62.77	21.58	37.76	51.65
Skewness	0.04	0.08	0.12	0.04	0.12	0.33
Kurtosis	0	0.14	0.68	0.71	4.58	9.16
Max. drawdown (%)	29.86	58.51	65.67	34.22	13.87	4.36
Sharpe ratio	0.66	0.2	0.07	0.11	0.09	0.15
Calmar ratio	0.55	0.21	0.1	0.12	0.05	0.09
Average stock allocation (%)	199.74	192.53	143.33	165.02	25.69	4.86
Qt. 1% average control	1.94	1.38	0.99	1.12	0.11	0.02
Qt. 50% average control	2	2	1.43	1.71	0.17	0.03
Qt. 99% average control	2	2	1.95	2	1.39	0.48

Table 3.10: Performance measurements for S-shape utility investor over 50,000 draws of portfolios. The protection level α is set to 0. The optimal control is calculated numerically. All amounts are simulated over 5 years and evaluated in terms of their discounted values. The corresponding terminal wealth distributions are presented in Figure 3.4.

S-shape, $\alpha = 0.8, \pi = p^*$	$\mu - r = 0.07$			$\mu - r = 0.02$		
	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$
Constant Equivalent	1.45	1.13	1.06	1.04	1	1
Mean terminal utility	0.5	0.17	0.08	0.05	0	0
Std. dev. terminal utility	0.57	0.48	0.39	0.3	0.1	0.04
Qt. 1% terminal utility	-0.45	-0.51	-0.51	-0.49	-0.51	-0.02
Qt. 50% terminal utility	0.41	0.08	0.01	0.01	0.01	0
Qt. 99% terminal utility	2.27	1.82	1.47	1.04	0.26	0.04
Mean Ann. return (%)	7.58	3.14	1.82	1.34	0.15	0.02
Min Ann. return (%)	-4.26	-4.27	-4.27	-4.26	-4.25	-4.27
Max Ann. return (%)	36.76	36.78	37.28	35.8	34.51	6
Std. dev. Ann. return	7.4	6.13	4.94	3.86	1.11	0.31
Mean Ann. volatility (%)	10.8	9.73	8.27	6.78	1.71	0.42
Min Ann. volatility (%)	2.55	2.14	2.03	1.98	0.71	0.23
Max Ann. volatility (%)	16.14	19.84	20.64	14.91	16.26	9.71
Skewness	0.35	0.37	0.3	0.19	0.16	0.39
Kurtosis	3.3	5.2	4.56	3.23	5.57	11.82
Max. drawdown (%)	15.28	17.01	16.29	14.37	4.25	1.06
Sharpe ratio	0.59	0.18	0.08	0.08	0.08	0.11
Calmar ratio	0.54	0.2	0.11	0.1	0.05	0.07
Average stock allocation (%)	96.51	39.89	22.92	59.16	7.06	1.13
Qt. 1% average control	9.45	2.97	1.5	2.99	0.18	0.04
Qt. 50% average control	9.94	6.96	3.7	6.19	0.26	0.04
Qt. 99% average control	9.99	8.74	7.47	9.64	3.99	0.45

Table 3.11: Performance measurements for S-shape utility investor over 50,000 draws of portfolios. The protection level α is set to 0.8. The optimal control is calculated numerically. All amounts are simulated over 5 years and evaluated in terms of their discounted values. The corresponding terminal wealth distributions are presented in Figure 3.5.

S-shape $\Theta = 1.05$, $\alpha = 0, \pi = p^*$	$\mu - r = 0.07$			$\mu - r = 0.02$		
	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$
Constant Equivalent	1.93	1.57	1.22	1.06	1.03	1.02
Mean terminal utility	0.89	0.57	0.21	0.01	-0.08	-0.1
Std. dev. terminal utility	0.88	1.72	1.76	0.67	0.4	0.39
Qt. 1% terminal utility	-1.03	-2.02	-2.26	-1.56	-1.87	-2.06
Qt. 50% terminal utility	0.8	0.1	0.03	0.02	0.01	-0.01
Qt. 99% terminal utility	3.46	5.34	5.35	1.83	0.37	0.05
Mean Ann. return (%)	13.2	7.71	1.98	2.12	0.14	-0.27
Min Ann. return (%)	-21.43	-48.43	-65.96	-28.91	-46.13	-59.71
Max Ann. return (%)	51.41	52.21	52.3	50.25	50.85	9.2
Std. dev. Ann. return	10.18	18.79	20.73	8.02	5.12	5.85
Mean Ann. volatility (%)	19.99	38.96	44.72	17.23	8.49	7.28
Min Ann. volatility (%)	16.05	19.84	20.1	11.82	2.52	0.97
Max Ann. volatility (%)	21.62	43.26	62.82	21.58	39.18	57.37
Skewness	0.04	0.08	0.12	0.05	0.27	0.64
Kurtosis	0	0.14	0.68	0.71	5.56	12.72
Max. drawdown (%)	29.86	58.5	65.44	34.17	17.63	14.14
Sharpe ratio	0.66	0.2	0.07	0.12	0.21	0.32
Calmar ratio	0.55	0.21	0.1	0.12	0.12	0.21
Average stock allocation (%)	199.71	192.47	143.09	164.77	34.53	17.98
Qt. 1% average control	1.94	1.39	0.99	1.13	0.11	0.03
Qt. 50% average control	2	2	1.42	1.7	0.2	0.07
Qt. 99% average control	2	2	1.96	2	1.67	1.48

Table 3.12: Performance measurements for S-shape utility investor with reference point $\Theta = 1.05$ over 50,000 draws of portfolios. The protection level α is set to 0. The optimal control is calculated numerically. All amounts are simulated over 5 years and evaluated in terms of their discounted values. The corresponding terminal wealth distributions are presented in Figure 3.6.

S-shape $\Theta = 1.05$, $\alpha = 0.8, \pi = p^*$	$\mu - r = 0.07$			$\mu - r = 0.02$		
	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$
Constant Equivalent	1.44	1.11	1.05	1.04	1.02	1.02
Mean terminal utility	0.43	0.08	-0.01	-0.04	-0.1	-0.11
Std. dev. terminal utility	0.59	0.5	0.41	0.32	0.21	0.22
Qt. 1% terminal utility	-0.57	-0.63	-0.63	-0.62	-0.63	-0.64
Qt. 50% terminal utility	0.36	0.02	0	0	-0.01	-0.01
Qt. 99% terminal utility	2.22	1.77	1.35	0.92	0.05	0.02
Mean Ann. return (%)	7.58	3.14	1.79	1.29	0.28	0.17
Min Ann. return (%)	-4.26	-4.28	-4.28	-4.26	-4.26	-4.27
Max Ann. return (%)	36.77	36.59	38.09	35.54	19.22	3
Std. dev. Ann. return	7.4	6.05	4.71	3.59	1.66	1.71
Mean Ann. volatility (%)	10.8	9.65	8.06	6.55	3.37	3.34
Min Ann. volatility (%)	2.54	2.13	2.01	1.95	0.93	0.59
Max Ann. volatility (%)	16.15	19.58	21.05	14.73	12.42	11.63
Skewness	0.35	0.35	0.29	0.18	0.36	0.48
Kurtosis	3.25	4.85	4.06	2.56	5.64	8.2
Max. drawdown (%)	15.27	16.94	16.04	13.99	7.88	7.89
Sharpe ratio	0.59	0.18	0.1	0.11	0.26	0.26
Calmar ratio	0.54	0.2	0.12	0.11	0.2	0.21
Average stock allocation (%)	96.51	40.08	22.86	58.31	13.97	8.92
Qt. 1% average control	9.5	3.25	1.48	2.95	0.23	0.08
Qt. 50% average control	9.92	6.79	3.39	5.82	0.65	0.41
Qt. 99% average control	9.98	8.67	7.14	9.38	6.25	5.54

Table 3.13: Performance measurements for S-shape utility investor with reference point $\Theta = 1.05$ over 50,000 draws of portfolios. The protection level α is set to 0.8. The optimal control is calculated numerically. All amounts are simulated over 5 years and evaluated in terms of their discounted values. The corresponding terminal wealth distributions are presented in Figure 3.7.

CRRA, $\alpha = 0, \pi = p^*$	$\mu - r = 0.07,\ \hat{\mu} - r = 0.02$			$\mu - r = 0.02,\ \hat{\mu} - r = 0.07$		
	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$
Constant Equivalent	1.34	1.08	1.03	1	0.94	0.97
Mean terminal utility	-0.74	-0.93	-0.97	-1	-1.06	-1.03
Std. dev. terminal utility	0.17	0.1	0.07	0.47	0.42	0.27
Qt. 1% terminal utility	-1.22	-1.19	-1.15	-2.58	-2.4	-1.8
Qt. 50% terminal utility	-0.73	-0.92	-0.96	-0.9	-0.99	-0.99
Qt. 99% terminal utility	-0.44	-0.72	-0.81	-0.32	-0.4	-0.55
Mean Ann. return (%)	6.71	1.64	0.73	2.45	0.56	0.25
Min Ann. return (%)	-10.64	-6.96	-5.05	-28.91	-25.75	-18.3
Max Ann. return (%)	29.29	11.79	7.32	50.24	40.97	25.17
Std. dev. Ann. return	4.72	2.24	1.48	9.21	7.72	5.14
Mean Ann. volatility (%)	9.82	4.9	3.27	20	17.09	11.41
Min Ann. volatility (%)	9.01	4.49	2.99	18.28	15.71	10.49
Max Ann. volatility (%)	10.62	5.3	3.53	21.63	18.5	12.35
Skewness	0.02	0.01	0.01	0.04	0.03	0.02
Kurtosis	0	0	0	0	0	0
Max. drawdown (%)	15.76	9.98	7.29	38.51	35.89	26.04
Sharpe ratio	0.68	0.33	0.22	0.12	0.03	0.02
Calmar ratio	0.54	0.25	0.18	0.13	0.08	0.07
Average stock allocation (%)	98.18	24.51	10.89	200	85.44	38.05
Qt. 1% average control	0.98	0.24	0.11	2	0.84	0.38
Qt. 50% average control	0.98	0.25	0.11	2	0.85	0.38
Qt. 99% average control	0.99	0.25	0.11	2	0.86	0.38

Table 3.14: Performance measurements for CRRA (2) investor over 50,000 draws of portfolios. The protection level α is set to 0. The optimal control is calculated numerically with mis-estimate of the risk premium $\hat{\mu} - r$. All amounts are simulated over 5 years and evaluated in terms of their discounted values. The corresponding terminal wealth distributions are presented in Figure 3.8.

CRRA, $\alpha = 0.8, \pi = p^*$	$\mu - r = 0.07,\ \hat{\mu} - r = 0.02$			$\mu - r = 0.02,\ \hat{\mu} - r = 0.07$		
	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$
Constant Equivalent	1.3	1.08	1.04	1.05	1.02	1.01
Mean terminal utility	-0.77	-0.92	-0.96	-0.95	-0.98	-0.99
Std. dev. terminal utility	0.16	0.11	0.07	0.19	0.18	0.16
Qt. 1% terminal utility	-1.16	-1.18	-1.15	-1.23	-1.23	-1.23
Qt. 50% terminal utility	-0.75	-0.92	-0.96	-0.98	-1.02	-1.02
Qt. 99% terminal utility	-0.45	-0.7	-0.8	-0.46	-0.51	-0.59
Mean Ann. return (%)	5.97	1.78	0.82	1.61	0.83	0.52
Min Ann. return (%)	-4.25	-4.26	-4.15	-4.26	-4.27	-4.28
Max Ann. return (%)	28.42	12.67	7.81	35.77	35.59	24.04
Std. dev. Ann. return	4.48	2.33	1.54	4.66	4.2	3.64
Mean Ann. volatility (%)	8.03	4.83	3.31	8.33	8.06	7.33
Min Ann. volatility (%)	2.58	2.54	2.36	2.1	2.01	1.93
Max Ann. volatility (%)	9.39	5.3	3.62	15.01	13.26	10.39
Skewness	0.16	0.08	0.06	0.25	0.19	0.15
Kurtosis	0.45	0.09	0.05	4.55	2.72	1.53
Max. drawdown (%)	12.47	9.75	7.34	16.49	17.07	16.47
Sharpe ratio	0.7	0.36	0.25	0.07	-0.02	-0.03
Calmar ratio	0.57	0.27	0.2	0.12	0.07	0.06
Average stock allocation (%)	78.32	23.97	10.98	71.28	35.99	22.79
Qt. 1% average control	5.16	1.38	0.6	6.67	3.18	2.01
Qt. 50% average control	5.72	1.55	0.66	9.52	5.42	3.01
Qt. 99% average control	8.18	3.64	1.17	9.85	8.18	6.78

Table 3.15: Performance measurements for CRRA (2) investor over 50,000 draws of portfolios. The protection level α is set to 0.8. The optimal control is calculated numerically with mis-estimate of the risk premium $\hat{\mu} - r$. All amounts are simulated over 5 years and evaluated in terms of their discounted values. The corresponding terminal wealth distributions are presented in Figure 3.9.

CRRA, $\alpha = 0, \pi = p^*$	$\mu - r = 0.07$			
	$\sigma = 0.1,\ \hat{\sigma} = 0.2$	$\sigma = 0.1,\ \hat{\sigma} = 0.3$	$\sigma = 0.3,\ \hat{\sigma} = 0.2$	$\sigma = 0.3,\ \hat{\sigma} = 0.1$
Constant Equivalent	1.3	1.13	0.97	0.33
Mean terminal utility	-0.77	-0.88	-1.03	-3.01
Std. dev. terminal utility	0.15	0.08	0.64	6.26
Qt. 1% terminal utility	-1.18	-1.07	-3.28	-27.75
Qt. 50% terminal utility	-0.75	-0.88	-0.87	-1.21
Qt. 99% terminal utility	-0.48	-0.72	-0.23	-0.12
Mean Ann. return (%)	5.88	2.65	3.44	-0.21
Min Ann. return (%)	-9.33	-4.2	-33.92	-60.62
Max Ann. return (%)	25.12	10.54	51.44	56.43
Std. dev. Ann. return	4.08	1.76	11.95	26.4
Mean Ann. volatility (%)	8.57	3.81	25.65	58.78
Min Ann. volatility (%)	7.87	3.49	20.17	23.47
Max Ann. volatility (%)	9.27	4.12	27.81	64.9
Skewness	0.02	0.01	0.05	0.13
Kurtosis	0	0	0	0.23
Max. drawdown (%)	13.89	6.38	45.94	76.6
Sharpe ratio	0.69	0.69	0.13	0.02
Calmar ratio	0.54	0.53	0.14	0.07
Average stock allocation (%)	85.73	38.11	85.46	193.36
Qt. 1% average control	0.85	0.38	0.83	0.81
Qt. 50% average control	0.86	0.38	0.86	2
Qt. 99% average control	0.86	0.38	0.87	2
CRRA, $\alpha = 0, \pi = p^*$	$\mu - r = 0.02$			
	$\sigma = 0.1,\ \hat{\sigma} = 0.2$	$\sigma = 0.1,\ \hat{\sigma} = 0.3$	$\sigma = 0.3,\ \hat{\sigma} = 0.2$	$\sigma = 0.3,\ \hat{\sigma} = 0.1$
Constant Equivalent	1.02	1.01	1	0.72
Mean terminal utility	-0.98	-0.99	-1	-1.39
Std. dev. terminal utility	0.05	0.02	0.17	0.99
Qt. 1% terminal utility	-1.11	-1.05	-1.45	-5
Qt. 50% terminal utility	-0.98	-0.99	-0.99	-1.13
Qt. 99% terminal utility	-0.86	-0.94	-0.67	-0.24
Mean Ann. return (%)	0.47	0.21	0.28	-1.44
Min Ann. return (%)	-3.87	-1.73	-12.12	-39.98
Max Ann. return (%)	5.36	2.35	15.66	52.59
Std. dev. Ann. return	1.1	0.49	3.31	12.97
Mean Ann. volatility (%)	2.45	1.09	7.35	29.26
Min Ann. volatility (%)	2.24	1	6.74	23.11
Max Ann. volatility (%)	2.65	1.18	7.94	31.77
Skewness	0	0	0.01	0.05
Kurtosis	0	0	0	0
Max. drawdown (%)	5.63	2.54	17.56	54.83
Sharpe ratio	0.19	0.2	0.04	-0.05
Calmar ratio	0.16	0.16	0.08	0.03
Average stock allocation (%)	24.5	10.89	24.49	97.52
Qt. 1% average control	0.24	0.11	0.24	0.94
Qt. 50% average control	0.25	0.11	0.24	0.98
Qt. 99% average control	0.25	0.11	0.25	0.99

Table 3.16: Performance measurements for CRRA (2) investor over 50,000 draws of portfolios. The protection level α is set to 0. The optimal control is calculated numerically with mis-estimate of the volatility $\hat{\sigma}$. All amounts are simulated over 5 years and evaluated in terms of their discounted values.

CRRA, $\alpha = 0.8, \pi = p^*$	$\mu - r = 0.07$			
	$\sigma = 0.1,\ \hat{\sigma} = 0.2$	$\sigma = 0.1,\ \hat{\sigma} = 0.3$	$\sigma = 0.3,\ \hat{\sigma} = 0.2$	$\sigma = 0.3,\ \hat{\sigma} = 0.1$
Constant Equivalent	1.23	1.12	1.06	1.02
Mean terminal utility	-0.81	-0.89	-0.94	-0.98
Std. dev. terminal utility	0.12	0.07	0.21	0.2
Qt. 1% terminal utility	-1.12	-1.06	-1.24	-1.24
Qt. 50% terminal utility	-0.8	-0.89	-0.99	-1.02
Qt. 99% terminal utility	-0.56	-0.74	-0.38	-0.4
Mean Ann. return (%)	4.54	2.44	1.94	1.11
Min Ann. return (%)	-3.95	-3.41	-4.28	-4.3
Max Ann. return (%)	20.26	9.97	36.62	40.04
Std. dev. Ann. return	3.22	1.61	5.59	5.2
Mean Ann. volatility (%)	6.16	3.32	9.36	8.96
Min Ann. volatility (%)	2.89	2.61	2.12	2.53
Max Ann. volatility (%)	7.04	3.64	18.64	25.62
Skewness	0.13	0.07	0.34	1.35
Kurtosis	0.51	0.19	6.75	55.99
Max. drawdown (%)	9.88	5.53	17.45	16.59
Sharpe ratio	0.71	0.73	0.05	-0.04
Calmar ratio	0.57	0.56	0.12	0.07
Average stock allocation (%)	60.09	32.87	24.28	11.6
Qt. 1% average control	3.68	1.82	1.58	0.53
Qt. 50% average control	3.88	1.88	6.37	9.86
Qt. 99% average control	5.08	2.23	8.97	9.97
CRRA, $\alpha = 0.8, \pi = p^*$	$\mu - r = 0.02$			
	$\sigma = 0.1,\ \hat{\sigma} = 0.2$	$\sigma = 0.1,\ \hat{\sigma} = 0.3$	$\sigma = 0.3,\ \hat{\sigma} = 0.2$	$\sigma = 0.3,\ \hat{\sigma} = 0.1$
Constant Equivalent	1.02	1.01	1.01	1
Mean terminal utility	-0.98	-0.99	-0.99	-1
Std. dev. terminal utility	0.06	0.02	0.15	0.18
Qt. 1% terminal utility	-1.11	-1.05	-1.23	-1.24
Qt. 50% terminal utility	-0.97	-0.99	-1	-1.05
Qt. 99% terminal utility	-0.85	-0.93	-0.66	-0.45
Mean Ann. return (%)	0.53	0.24	0.46	0.48
Min Ann. return (%)	-3.6	-1.74	-4.27	-4.29
Max Ann. return (%)	5.64	2.47	17.06	37.28
Std. dev. Ann. return	1.14	0.5	3.13	4.53
Mean Ann. volatility (%)	2.45	1.1	6.49	8.43
Min Ann. volatility (%)	2.16	1	1.89	2.29
Max Ann. volatility (%)	2.66	1.21	7.91	23.2
Skewness	0.04	0.03	0.11	0.59
Kurtosis	0.04	0.03	0.78	22.36
Max. drawdown (%)	5.62	2.56	15.07	17.25
Sharpe ratio	0.21	0.21	0.01	-0.09
Calmar ratio	0.17	0.17	0.08	0.03
Average stock allocation (%)	24.42	11	20.74	14.87
Qt. 1% average control	1.32	0.57	1.47	0.8
Qt. 50% average control	1.4	0.59	2.14	9.15
Qt. 99% average control	1.92	0.66	6.93	9.83

Table 3.17: Performance measurements for CRRA (2) investor over 50,000 draws of portfolios. The protection level α is set to 0.8. The optimal control is calculated numerically with mis-estimate of the volatility $\hat{\sigma}$. All amounts are simulated over 5 years and evaluated in terms of their discounted values.

S-shape, $\alpha = 0, \pi = p^*$	$\mu - r = 0.07,\ \hat{\mu} - r = 0.02$			$\mu - r = 0.02,\ \hat{\mu} - r = 0.07$		
	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$
Constant Equivalent	1.85	1.08	1.01	1.05	0.94	0.92
Mean terminal utility	0.86	0.11	0.02	0.07	-0.18	-0.26
Std. dev. terminal utility	0.86	0.34	0.09	0.75	1.38	1.5
Qt. 1% terminal utility	-0.79	-0.11	-0.03	-1.46	-2.05	-2.19
Qt. 50% terminal utility	0.73	0.04	0.02	0.11	-0.01	0.01
Qt. 99% terminal utility	3.47	1.58	0.17	1.94	5.24	5.33
Mean Ann. return (%)	11.93	1.65	0.26	2.44	-2.17	-5.29
Min Ann. return (%)	-21.43	-34.71	-43.02	-28.91	-53.34	-60.67
Max Ann. return (%)	51.34	51.01	22.29	50.27	51.89	52.25
Std. dev. Ann. return	10.42	4.34	1.22	9.19	17.31	20.27
Mean Ann. volatility (%)	18.29	6.43	1.58	19.96	39.14	45.94
Min Ann. volatility (%)	11.81	2.42	0.76	18.27	21.44	22
Max Ann. volatility (%)	21.42	37.24	49.55	21.63	43.24	62.93
Skewness	0.03	0.06	0.23	0.04	0.08	0.12
Kurtosis	0.39	4.03	6.9	0.01	0.11	0.63
Max. drawdown (%)	27.93	13.1	3.51	38.45	64.91	70.09
Sharpe ratio	0.61	0.19	0.23	0.12	-0.05	-0.08
Calmar ratio	0.52	0.12	0.14	0.13	0.02	-0.02
Average stock allocation (%)	178.99	27.17	4.24	199.43	193.88	147.59
Qt. 1% average control	1.15	0.11	0.02	1.94	1.71	1.01
Qt. 50% average control	1.93	0.19	0.03	2	2	1.49
Qt. 99% average control	2	1.2	0.22	2	2	1.96

Table 3.18: Performance measurements for S-shape utility investor over 50,000 draws of portfolios. The protection level α is set to 0. The optimal control is calculated numerically with mis-estimate of the risk premium $\hat{\mu} - r$. All amounts are simulated over 5 years and evaluated in terms of their discounted values. The corresponding terminal wealth distributions are presented in Figure 3.10.

S-shape, $\alpha = 0.8, \pi = p^*$	$\mu - r = 0.07,\ \hat{\mu} - r = 0.02$			$\mu - r = 0.02,\ \hat{\mu} - r = 0.07$		
	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$
Constant Equivalent	1.39	1.03	1	1.02	0.99	0.99
Mean terminal utility	0.43	0.04	0.01	0.04	-0.02	-0.03
Std. dev. terminal utility	0.49	0.13	0.03	0.38	0.38	0.33
Qt. 1% terminal utility	-0.37	-0.02	-0.02	-0.51	-0.53	-0.52
Qt. 50% terminal utility	0.32	0.02	0.01	0.01	-0.01	0
Qt. 99% terminal utility	2.04	0.62	0.06	1.18	1.19	1.06
Mean Ann. return (%)	6.46	0.65	0.08	1.58	0.81	0.53
Min Ann. return (%)	-4.26	-4.25	-4.25	-4.26	-4.27	-4.28
Max Ann. return (%)	36.57	35.73	10.61	35.83	36.64	36.72
Std. dev. Ann. return	6.72	1.84	0.25	4.73	4.53	3.96
Mean Ann. volatility (%)	9.15	2.07	0.38	8.4	8.33	7.43
Min Ann. volatility (%)	2.81	0.71	0.23	2.11	2.05	2
Max Ann. volatility (%)	15.71	16.51	8.71	15.81	19.48	20.62
Skewness	0.27	0.1	0.23	0.31	0.3	0.26
Kurtosis	2.66	4.58	7.35	6.92	6.51	4.76
Max. drawdown (%)	13.73	4.44	0.9	16.53	17.28	16.51
Sharpe ratio	0.56	0.19	0.2	0.06	-0.04	-0.06
Calmar ratio	0.49	0.11	0.13	0.11	0.05	0.04
Average stock allocation (%)	82.46	8.52	1.1	68.88	32.71	20.43
Qt. 1% average control	3.04	0.18	0.04	6.29	2.59	1.5
Qt. 50% average control	8.15	0.29	0.04	9.95	6.99	3.61
Qt. 99% average control	9.68	4.35	0.26	9.99	8.88	7.57

Table 3.19: Performance measurements for S-shape utility investor over 50,000 draws of portfolios. The protection level α is set to 0.8. The optimal control is calculated numerically with mis-estimate of the risk premium $\hat{\mu} - r$. All amounts are simulated over 5 years and evaluated in terms of their discounted values. The corresponding terminal wealth distributions are presented in Figure 3.11.

S-shape, $\alpha = 0, \pi = p^*$	$\mu - r = 0.07$			
	$\sigma = 0.1,\ \hat{\sigma} = 0.2$	$\sigma = 0.1,\ \hat{\sigma} = 0.3$	$\sigma = 0.3,\ \hat{\sigma} = 0.2$	$\sigma = 0.3,\ \hat{\sigma} = 0.1$
Constant Equivalent	1.92	1.61	1.2	1.17
Mean terminal utility	0.93	0.65	0.24	0.21
Std. dev. terminal utility	0.86	0.64	2.2	2.26
Qt. 1% terminal utility	-0.91	-0.57	-2.18	-2.18
Qt. 50% terminal utility	0.84	0.51	0	-0.38
Qt. 99% terminal utility	3.5	2.61	5.51	5.78
Mean Ann. return (%)	12.99	9.39	-0.04	-0.21
Min Ann. return (%)	-21.43	-18.16	-60.62	-60.62
Max Ann. return (%)	50.55	49.92	53.5	57.2
Std. dev. Ann. return	10.29	8.24	25.95	26.38
Mean Ann. volatility (%)	19.71	14.21	57.67	58.76
Min Ann. volatility (%)	16.34	10.23	22.06	22.81
Max Ann. volatility (%)	21.62	18.79	64.86	64.9
Skewness	0.04	0.02	0.13	0.13
Kurtosis	0.06	0.47	0.31	0.23
Max. drawdown (%)	29.58	22.56	75.82	76.59
Sharpe ratio	0.65	0.61	0.03	0.02
Calmar ratio	0.55	0.49	0.07	0.07
Average stock allocation (%)	196.12	138.88	188.43	193.22
Qt. 1% average control	1.71	0.99	0.85	0.82
Qt. 50% average control	2	1.43	1.99	2
Qt. 99% average control	2	1.69	2	2
S-shape, $\alpha = 0, \pi = p^*$	$\mu - r = 0.02$			
	$\sigma = 0.1,\ \hat{\sigma} = 0.2$	$\sigma = 0.1,\ \hat{\sigma} = 0.3$	$\sigma = 0.3,\ \hat{\sigma} = 0.2$	$\sigma = 0.3,\ \hat{\sigma} = 0.1$
Constant Equivalent	1.01	1	1	0.87
Mean terminal utility	0.02	0	-0.02	-0.37
Std. dev. terminal utility	0.1	0.02	0.52	1.78
Qt. 1% terminal utility	-0.51	-0.04	-2.03	-2.21
Qt. 50% terminal utility	0.02	0.01	0.01	-0.03
Qt. 99% terminal utility	0.2	0.03	1.33	5.61
Mean Ann. return (%)	0.32	0.05	-0.41	-8.11
Min Ann. return (%)	-17.53	-6.29	-60.58	-61.03
Max Ann. return (%)	17.31	0.91	52.47	56.39
Std. dev. Ann. return	1.07	0.16	7.43	23.44
Mean Ann. volatility (%)	2.04	0.31	11.97	55.18
Min Ann. volatility (%)	1.14	0.26	3.65	24.56
Max Ann. volatility (%)	16.16	8.74	59.34	64.71
Skewness	0.08	0.05	0.19	0.13
Kurtosis	3.42	3.25	6.75	0.53
Max. drawdown (%)	4.8	0.74	23.94	77.73
Sharpe ratio	0.21	0.19	0.03	-0.12
Calmar ratio	0.13	0.15	0.02	-0.05
Average stock allocation (%)	17.41	2.73	31.95	177
Qt. 1% average control	0.11	0.03	0.11	0.98
Qt. 50% average control	0.13	0.03	0.19	1.9
Qt. 99% average control	0.77	0.05	1.66	2

Table 3.20: Performance measurements for S-shape utility investor over 50,000 draws of portfolios. The protection level α is set to 0. The optimal control is calculated numerically with mis-estimate of the volatility $\hat{\sigma}$. All amounts are simulated over 5 years and evaluated in terms of their discounted values.

S-shape, $\alpha = 0.8, \pi = p^*$	$\mu - r = 0.07$			
	$\sigma = 0.1,\ \hat{\sigma} = 0.2$	$\sigma = 0.1,\ \hat{\sigma} = 0.3$	$\sigma = 0.3,\ \hat{\sigma} = 0.2$	$\sigma = 0.3,\ \hat{\sigma} = 0.1$
Constant Equivalent	1.36	1.18	1.02	1
Mean terminal utility	0.4	0.22	0.04	-0.02
Std. dev. terminal utility	0.39	0.23	0.45	0.43
Qt. 1% terminal utility	-0.33	-0.18	-0.53	-0.54
Qt. 50% terminal utility	0.35	0.16	0.01	-0.09
Qt. 99% terminal utility	1.52	0.95	1.58	1.42
Mean Ann. return (%)	6.18	3.41	1.64	1.02
Min Ann. return (%)	-4.06	-3.54	-4.29	-4.3
Max Ann. return (%)	34.79	24.18	37.77	40.59
Std. dev. Ann. return	5.49	3.49	5.5	5.13
Mean Ann. volatility (%)	8.51	4.88	9.26	8.95
Min Ann. volatility (%)	2.8	2.01	2.26	2.63
Max Ann. volatility (%)	12.06	8.93	25.2	25.49
Skewness	0.22	0.15	0.59	1.69
Kurtosis	1.6	2.45	14.87	79.07
Max. drawdown (%)	13.02	8.33	17.19	16.35
Sharpe ratio	0.63	0.56	0.02	-0.05
Calmar ratio	0.52	0.41	0.1	0.07
Average stock allocation (%)	79.88	44.72	19.56	10.28
Qt. 1% average control	3.96	1.43	1.19	0.45
Qt. 50% average control	6.54	2.71	7.81	9.98
Qt. 99% average control	7.28	4.55	9.42	10
S-shape, $\alpha = 0.8, \pi = p^*$	$\mu - r = 0.02$			
	$\sigma = 0.1,\ \hat{\sigma} = 0.2$	$\sigma = 0.1,\ \hat{\sigma} = 0.3$	$\sigma = 0.3,\ \hat{\sigma} = 0.2$	$\sigma = 0.3,\ \hat{\sigma} = 0.1$
Constant Equivalent	1	1	1	0.98
Mean terminal utility	0.01	0	0	-0.06
Std. dev. terminal utility	0.03	0.01	0.16	0.38
Qt. 1% terminal utility	-0.02	-0.02	-0.52	-0.53
Qt. 50% terminal utility	0.01	0	0.01	-0.12
Qt. 99% terminal utility	0.07	0.01	0.53	1.17
Mean Ann. return (%)	0.1	0.02	0.19	0.48
Min Ann. return (%)	-4	-1.13	-4.28	-4.29
Max Ann. return (%)	7.93	0.44	35.91	38.45
Std. dev. Ann. return	0.32	0.04	1.86	4.5
Mean Ann. volatility (%)	0.58	0.09	2.96	8.42
Min Ann. volatility (%)	0.35	0.08	1.09	2.31
Max Ann. volatility (%)	5.26	2.38	21.04	25.46
Skewness	0.1	0.03	0.22	1.03
Kurtosis	3.76	1.29	8.26	36.72
Max. drawdown (%)	1.39	0.2	7.17	16.67
Sharpe ratio	0.2	0.2	0.02	-0.09
Calmar ratio	0.12	0.16	0.03	0.04
Average stock allocation (%)	4.98	0.81	7.57	12.54
Qt. 1% average control	0.18	0.04	0.19	0.67
Qt. 50% average control	0.22	0.04	0.32	9.34
Qt. 99% average control	1.14	0.05	6.38	9.94

Table 3.21: Performance measurements for S-shape utility investor over 50,000 draws of portfolios. The protection level α is set to 0.8. The optimal control is calculated numerically with mis-estimate of the volatility $\hat{\sigma}$. All amounts are simulated over 5 years and evaluated in terms of their discounted values.

CRRA (1Y), $\alpha = 0, \pi = p^*$	$\mu - r = 0.07$			$\mu - r = 0.02$		
	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$
Constant Equivalent	1.11	1.03	1.01	1.01	1	1
Mean terminal utility	-0.9	-0.97	-0.99	-0.99	-1	-1
Std. dev. terminal utility	0.18	0.17	0.11	0.1	0.05	0.03
Qt. 1% terminal utility	-1.42	-1.43	-1.28	-1.24	-1.12	-1.08
Qt. 50% terminal utility	-0.89	-0.95	-0.98	-0.98	-1	-1
Qt. 99% terminal utility	-0.56	-0.64	-0.75	-0.78	-0.89	-0.93
Mean Ann. return (%)	15.13	6.27	2.76	2.03	0.51	0.23
Min Ann. return (%)	-50.71	-48.27	-36.24	-32.21	-18.02	-12.48
Max Ann. return (%)	155.64	111.97	63.24	51.81	22.71	14.51
Std. dev. Ann. return	23.23	18.36	11.79	10.03	4.93	3.28
Mean Ann. volatility (%)	19.99	17.14	11.42	9.79	4.9	3.26
Min Ann. volatility (%)	16.27	13.91	9.28	7.95	3.99	2.65
Max Ann. volatility (%)	23.61	20.19	13.45	11.55	5.79	3.86
Skewness	0.04	0.03	0.02	0.02	0.01	0.01
Kurtosis	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02
Max. drawdown (%)	17.28	16.63	11.81	10.31	5.46	3.71
Sharpe ratio	0.76	0.36	0.24	0.21	0.1	0.07
Calmar ratio	1.41	0.8	0.63	0.58	0.45	0.41
Average stock allocation (%)	200	85.73	38.11	97.99	24.5	10.89
Qt. 1% average control	2	0.85	0.38	0.98	0.24	0.11
Qt. 50% average control	2	0.86	0.38	0.98	0.25	0.11
Qt. 99% average control	2	0.86	0.38	0.98	0.25	0.11

Table 3.22: Performance measurements for CRRA (2) investor over 50,000 draws of portfolios. The protection level α is set to 0. The optimal control is calculated numerically. All amounts are simulated over 1 year and evaluated in terms of their discounted values. The corresponding terminal wealth distributions are presented in Figure 3.12.

CRRA (1Y), $\alpha = 0.8, \pi = p^*$	$\mu - r = 0.07$			$\mu - r = 0.02$		
	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$
Constant Equivalent	1.09	1.04	1.02	1.02	1.01	1
Mean terminal utility	-0.92	-0.96	-0.98	-0.98	-0.99	-1
Std. dev. terminal utility	0.14	0.14	0.11	0.1	0.05	0.03
Qt. 1% terminal utility	-1.18	-1.21	-1.21	-1.18	-1.12	-1.08
Qt. 50% terminal utility	-0.92	-0.96	-0.97	-0.98	-0.99	-1
Qt. 99% terminal utility	-0.6	-0.64	-0.73	-0.77	-0.88	-0.92
Mean Ann. return (%)	11.73	6.76	3.65	2.75	0.82	0.41
Min Ann. return (%)	-19.46	-19.58	-19.53	-19.18	-17.92	-12.87
Max Ann. return (%)	137.48	116.71	69.79	57.11	24.66	15.52
Std. dev. Ann. return	18.28	17.13	12.15	10.28	5.17	3.41
Mean Ann. volatility (%)	13.48	14.04	10.98	9.43	4.99	3.32
Min Ann. volatility (%)	4.7	4.67	5.11	4.75	4.01	2.68
Max Ann. volatility (%)	19.58	19.04	13.57	11.72	5.93	3.91
Skewness	0.27	0.24	0.19	0.15	0.09	0.07
Kurtosis	1.24	0.72	0.24	0.21	0.06	0.05
Max. drawdown (%)	11.23	13.21	11.16	9.74	5.51	3.75
Sharpe ratio	0.7	0.33	0.27	0.23	0.16	0.12
Calmar ratio	1.42	0.85	0.72	0.66	0.53	0.48
Average stock allocation (%)	127.37	67.35	35.9	92.7	24.73	11
Qt. 1% average control	10	4.88	2.22	5.63	1.35	0.59
Qt. 50% average control	10	6.01	2.66	6.49	1.45	0.62
Qt. 99% average control	10	8.59	6.02	8.96	2.08	0.75

Table 3.23: Performance measurements for CRRA (2) investor over 50,000 draws of portfolios. The protection level α is set to 0.8. The optimal control is calculated numerically. All amounts are simulated over 1 year and evaluated in terms of their discounted values. The corresponding terminal wealth distributions are presented in Figure 3.13.

S-shape (1Y), $\alpha = 0, \pi = p^*$	$\mu - r = 0.07$			$\mu - r = 0.02$		
	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$
Constant Equivalent	1.09	1.02	1	1	1	1
Mean terminal utility	0.13	0.04	0.01	0	0	0
Std. dev. terminal utility	0.33	0.44	0.26	0.08	0.03	0.01
Qt. 1% terminal utility	-0.77	-1.29	-1.3	-0.42	-0.03	-0.03
Qt. 50% terminal utility	0.14	0.02	0.02	0.01	0	0
Qt. 99% terminal utility	0.82	1.3	0.69	0.13	0.02	0.01
Mean Ann. return (%)	14.69	8.66	2.4	0.42	0.04	0.02
Min Ann. return (%)	-50.71	-78.01	-85.89	-45.25	-48.18	-31.98
Max Ann. return (%)	155.64	393.25	400.23	68.4	6.5	7.61
Std. dev. Ann. return	23.09	31.95	16.66	3.59	0.95	0.35
Mean Ann. volatility (%)	19.57	25.18	12.52	2.7	0.56	0.31
Min Ann. volatility (%)	14.37	12.48	4.49	1.1	0.38	0.24
Max Ann. volatility (%)	23.55	45.45	61.34	20.21	32.57	32.16
Skewness	0.04	0.09	0.22	0.34	0.16	0.01
Kurtosis	0.08	1.63	4.31	6.99	3.22	0.52
Max. drawdown (%)	16.97	23.39	12.31	2.9	0.62	0.36
Sharpe ratio	0.74	0.29	0.22	0.2	0.13	0.07
Calmar ratio	1.38	0.59	0.33	0.31	0.42	0.41
Average stock allocation (%)	193.96	115.96	34.85	22.09	2.48	1
Qt. 1% average control	1.68	0.64	0.15	0.13	0.02	0.01
Qt. 50% average control	2	1.04	0.23	0.15	0.02	0.01
Qt. 99% average control	2	1.95	1.55	1.34	0.04	0.01

Table 3.24: Performance measurements for S-shape utility investor investor over 50,000 draws of portfolios. The protection level α is set to 0. The optimal control is calculated numerically. All amounts are simulated over 1 year and evaluated in terms of their discounted values. The corresponding terminal wealth distributions are presented in Figure 3.14.

S-shape (1Y), $\alpha = 0.8, \pi = p^*$	$\mu - r = 0.07$			$\mu - r = 0.02$		
	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$
Constant Equivalent	1.07	1.02	1	1	1	1
Mean terminal utility	0.1	0.03	0.01	0	0	0
Std. dev. terminal utility	0.25	0.23	0.11	0.05	0.01	0.01
Qt. 1% terminal utility	-0.42	-0.48	-0.49	-0.3	-0.02	-0.01
Qt. 50% terminal utility	0.1	0.01	0.01	0	0	0
Qt. 99% terminal utility	0.7	0.81	0.35	0.1	0.01	0.01
Mean Ann. return (%)	11.48	5.12	1.2	0.35	0.04	0.01
Min Ann. return (%)	-19.46	-19.53	-19.51	-18.38	-18.75	-15.85
Max Ann. return (%)	137.44	245.23	212.3	54.68	7.9	3.41
Std. dev. Ann. return	18.23	17.18	7.09	2.41	0.4	0.16
Mean Ann. volatility (%)	13.27	11.12	4.32	1.65	0.27	0.14
Min Ann. volatility (%)	4.8	4.78	1.54	0.64	0.18	0.11
Max Ann. volatility (%)	19.56	30.81	33.5	14.96	13.88	8.27
Skewness	0.27	0.27	0.26	0.38	0.17	0.03
Kurtosis	1.36	3.3	4.98	6.79	2.75	0.26
Max. drawdown (%)	11.12	10.71	4.51	1.78	0.29	0.16
Sharpe ratio	0.68	0.19	0.19	0.21	0.15	0.09
Calmar ratio	1.39	0.49	0.3	0.32	0.45	0.43
Average stock allocation (%)	124.34	47.58	11.72	13.35	1.2	0.46
Qt. 1% average control	8.14	1.68	0.3	0.38	0.06	0.02
Qt. 50% average control	10	3.14	0.45	0.47	0.06	0.02
Qt. 99% average control	10	9.35	5.59	5.64	0.1	0.02

Table 3.25: Performance measurements for S-shape utility investor over 50,000 draws of portfolios. The protection level α is set to 0.8. The optimal control is calculated numerically. All amounts are simulated over 1 year and evaluated in terms of their discounted values. The corresponding terminal wealth distributions are presented in Figure 3.15.

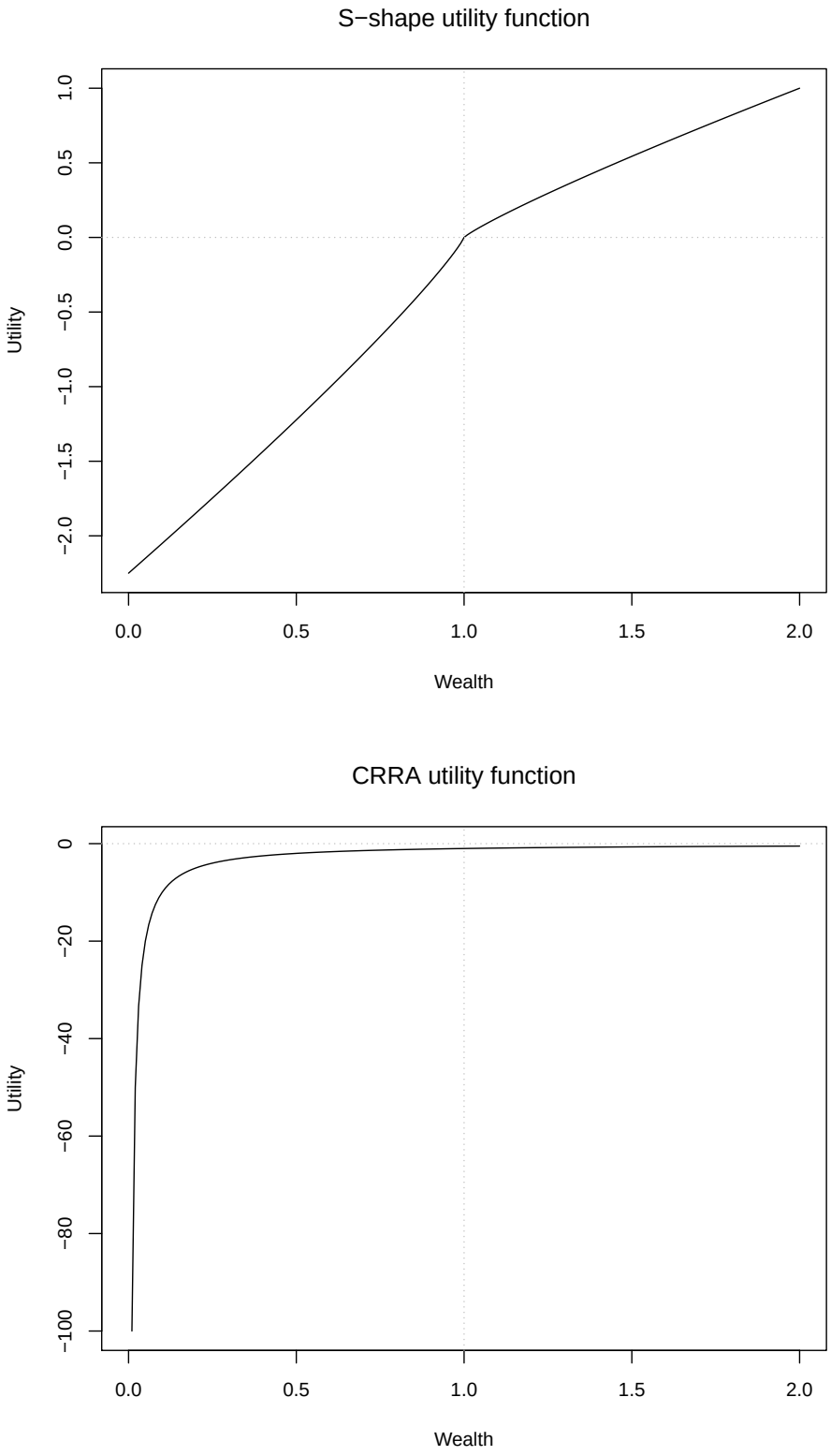

Figure 3.1: S-shape and CRRA utility functions. The dotted grey lines correspond to the reference point Θ of the S-shape utility (set to the initial wealth equal to 1). In the upper panel, the parameters are set as follows: $A = 2.25, B = 1, \gamma_1 = \gamma_2 = 0.88$. In the bottom panel γ is equal to 2.

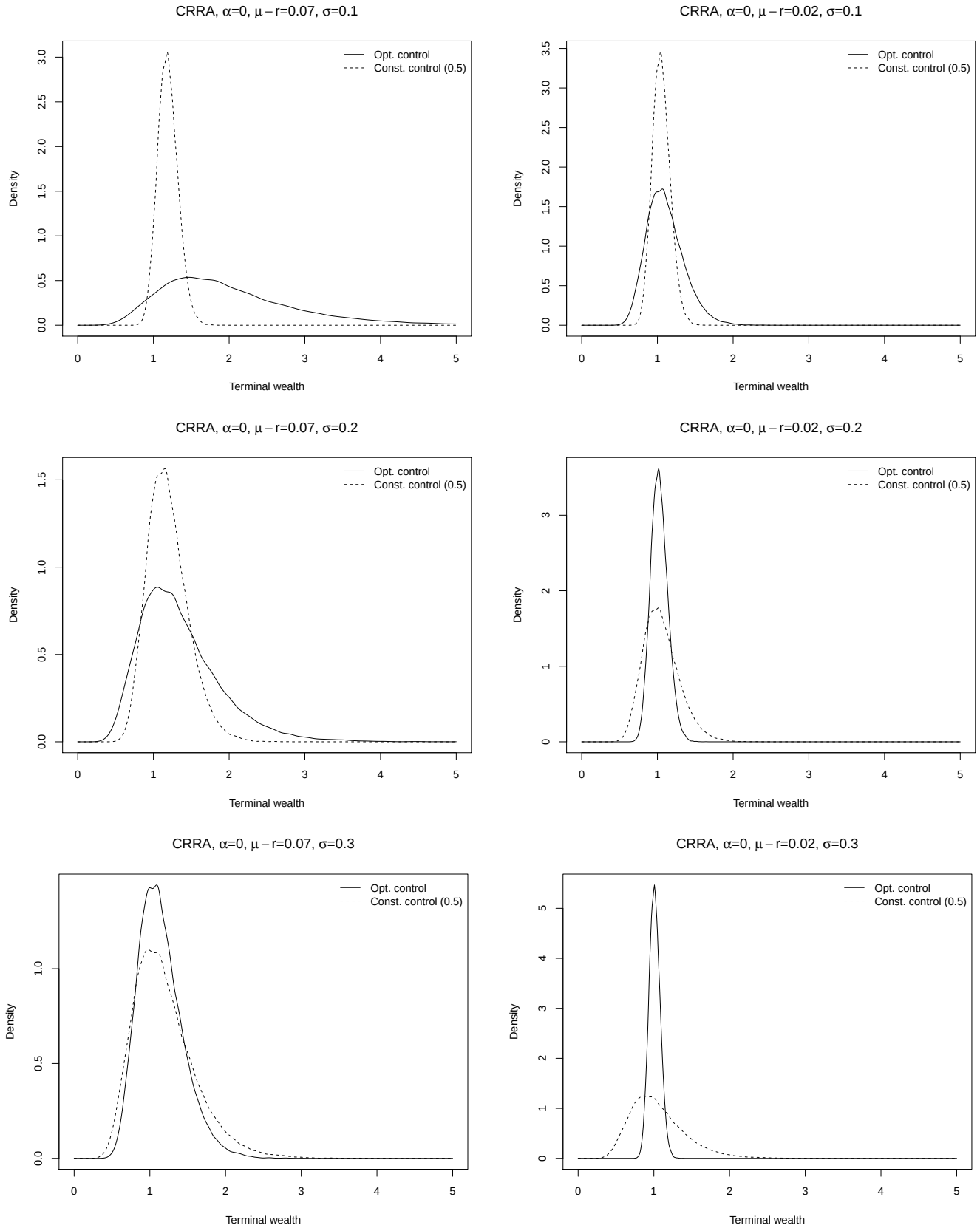

Figure 3.2: Terminal wealth distributions for the CRRA (2) investor with optimal control and constant control (equal to 0.5), without capital protection $\alpha = 0$. The different panels correspond to different parameter sets. All amounts are simulated over 5 years and evaluated in terms of their discounted values. Distributions are based on 50,000 draws for each configuration.

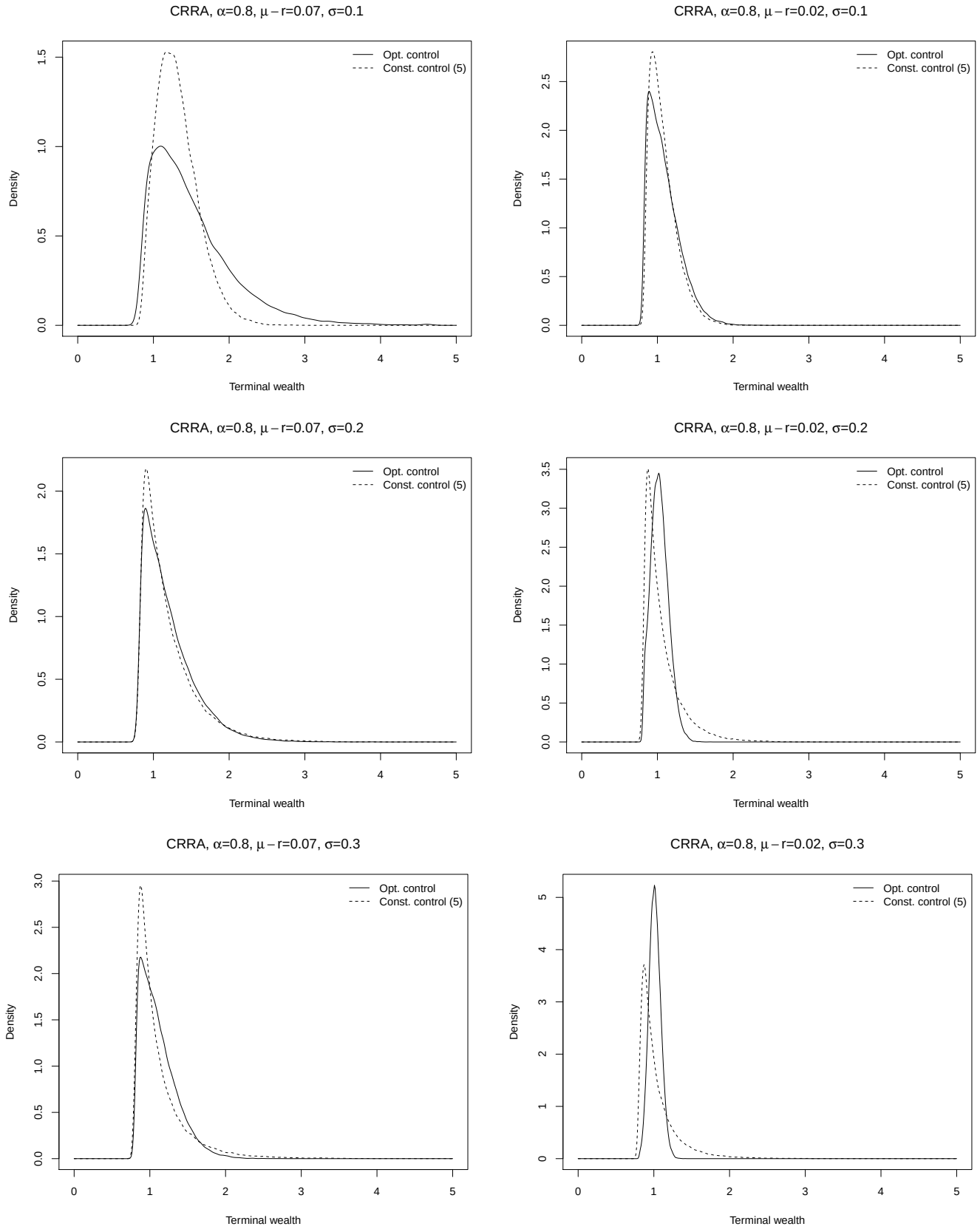

Figure 3.3: Terminal wealth distributions for the CRRA (2) investor with optimal control and constant control (equal to 5), with capital protection $\alpha = 0.8$. The different panels correspond to different parameter sets. All amounts are simulated over 5 years and evaluated in terms of their discounted values. Distributions are based on 50,000 draws for each configuration.

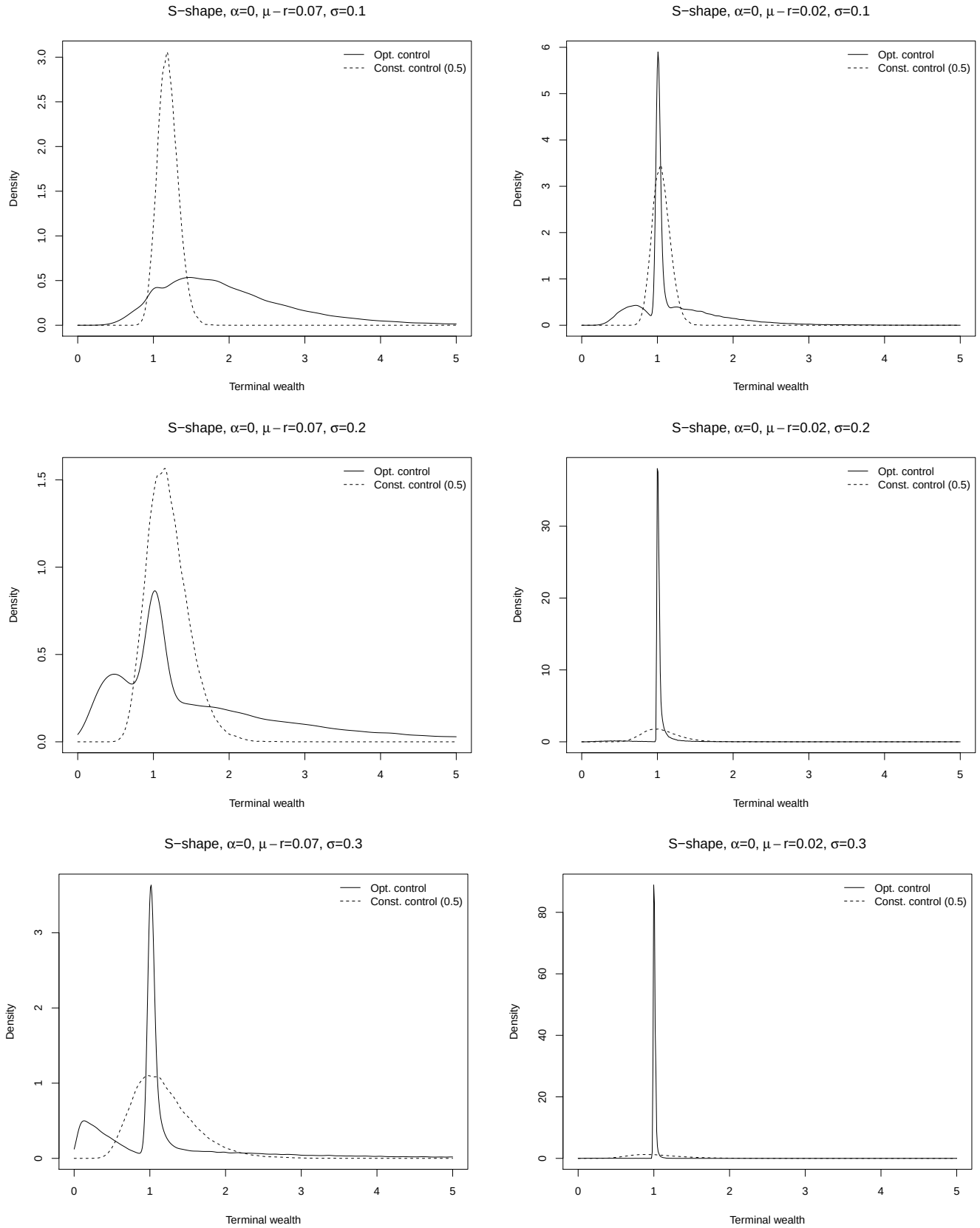

Figure 3.4: Terminal wealth distributions for the investor with S-shape utility with optimal control and constant control (equal to 0.5), without capital protection $\alpha = 0$. The different panels correspond to different parameter sets. All amounts are simulated over 5 years and evaluated in terms of their discounted values. Distributions are based on 50,000 draws for each configuration.

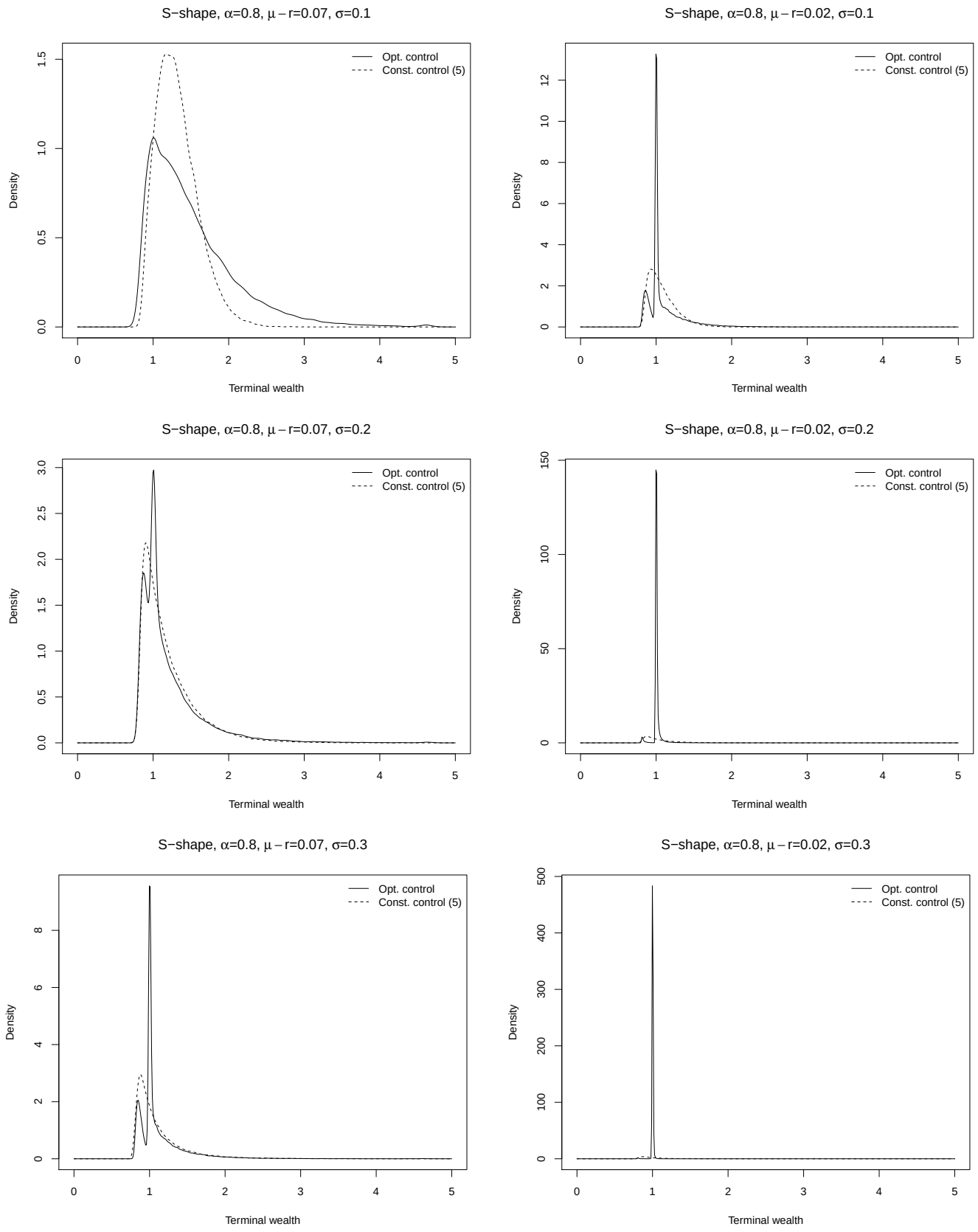

Figure 3.5: Terminal wealth distributions for the investor with S-shape utility with optimal control and constant control (equal to 5), with capital protection $\alpha = 0.8$. The different panels correspond to different parameter sets. All amounts are simulated over 5 years and evaluated in terms of their discounted values. Distributions are based on 50,000 draws for each configuration.

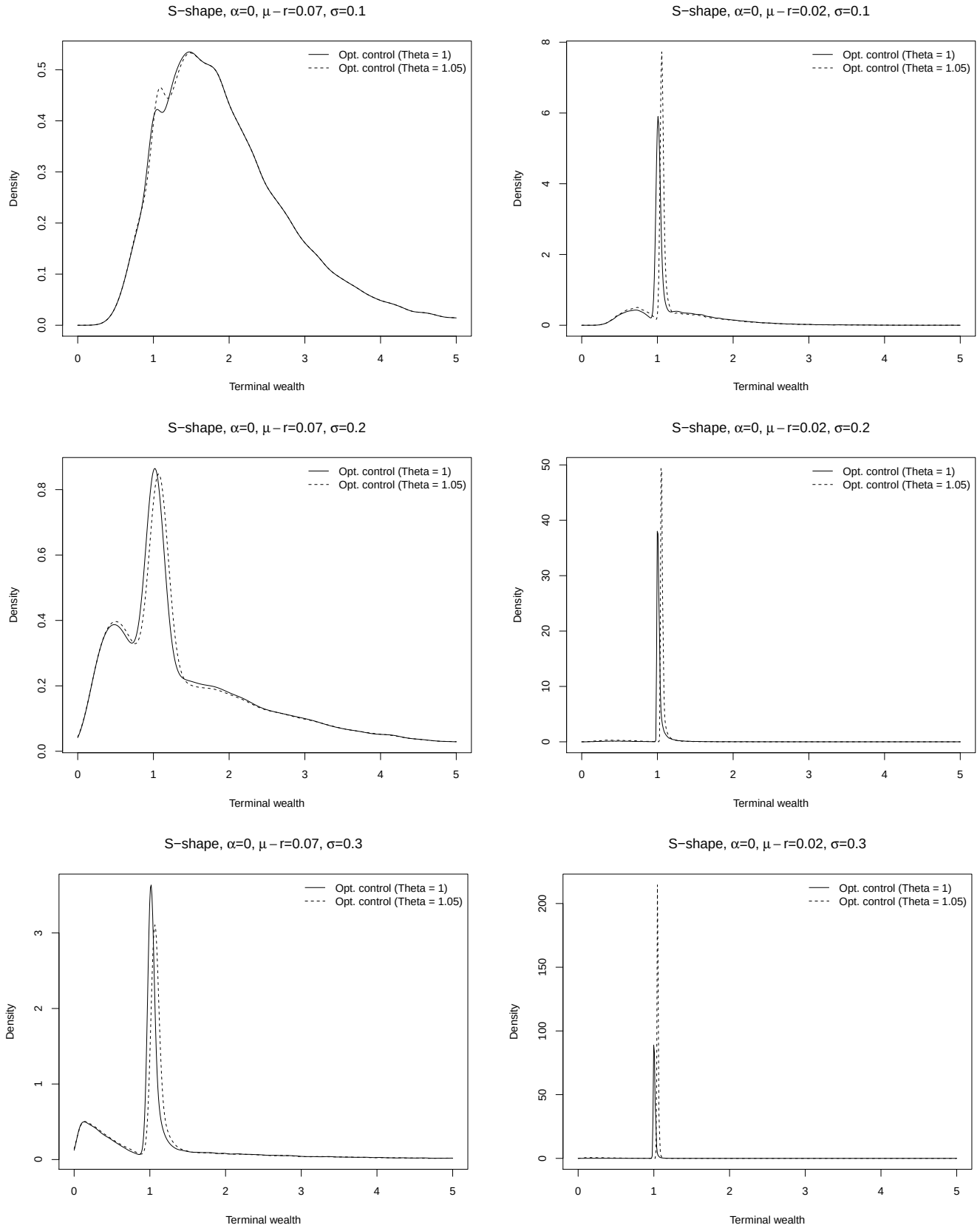

Figure 3.6: Terminal wealth distributions for the investor with S-shape utility with optimal control when $\Theta = 1$ and $\Theta = 1.05$, without capital protection $\alpha = 0$. The different panels correspond to different parameter sets. All amounts are simulated over 5 years and evaluated in terms of their discounted values. Distributions are based on 50,000 draws for each configuration.

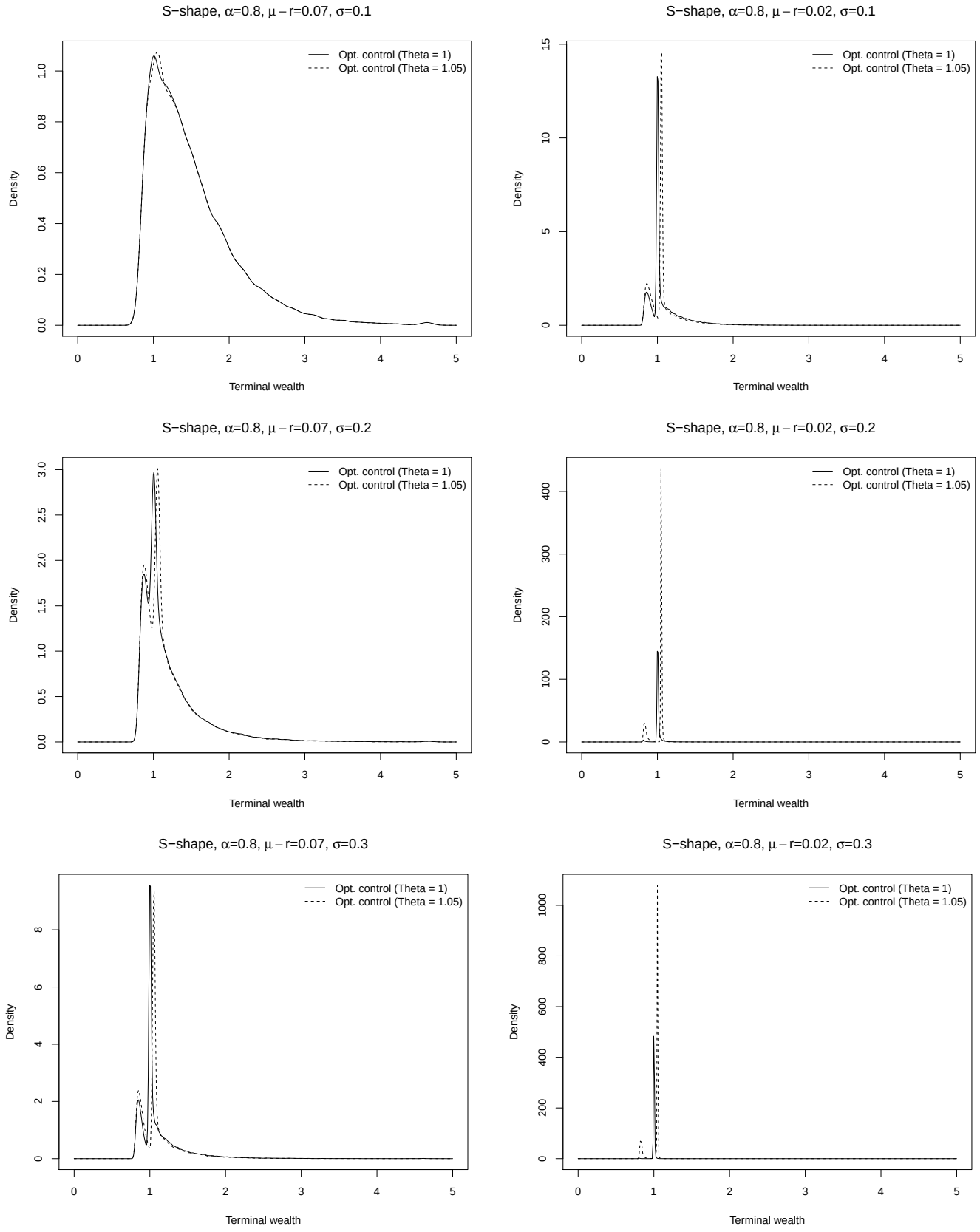

Figure 3.7: Terminal wealth distributions for the investor with S-shape utility with optimal control when $\Theta = 1$ and $\Theta = 1.05$, with capital protection $\alpha = 0.8$. The different panels correspond to different parameter sets. All amounts are simulated over 5 years and evaluated in terms of their discounted values. Distributions are based on 50,000 draws for each configuration.

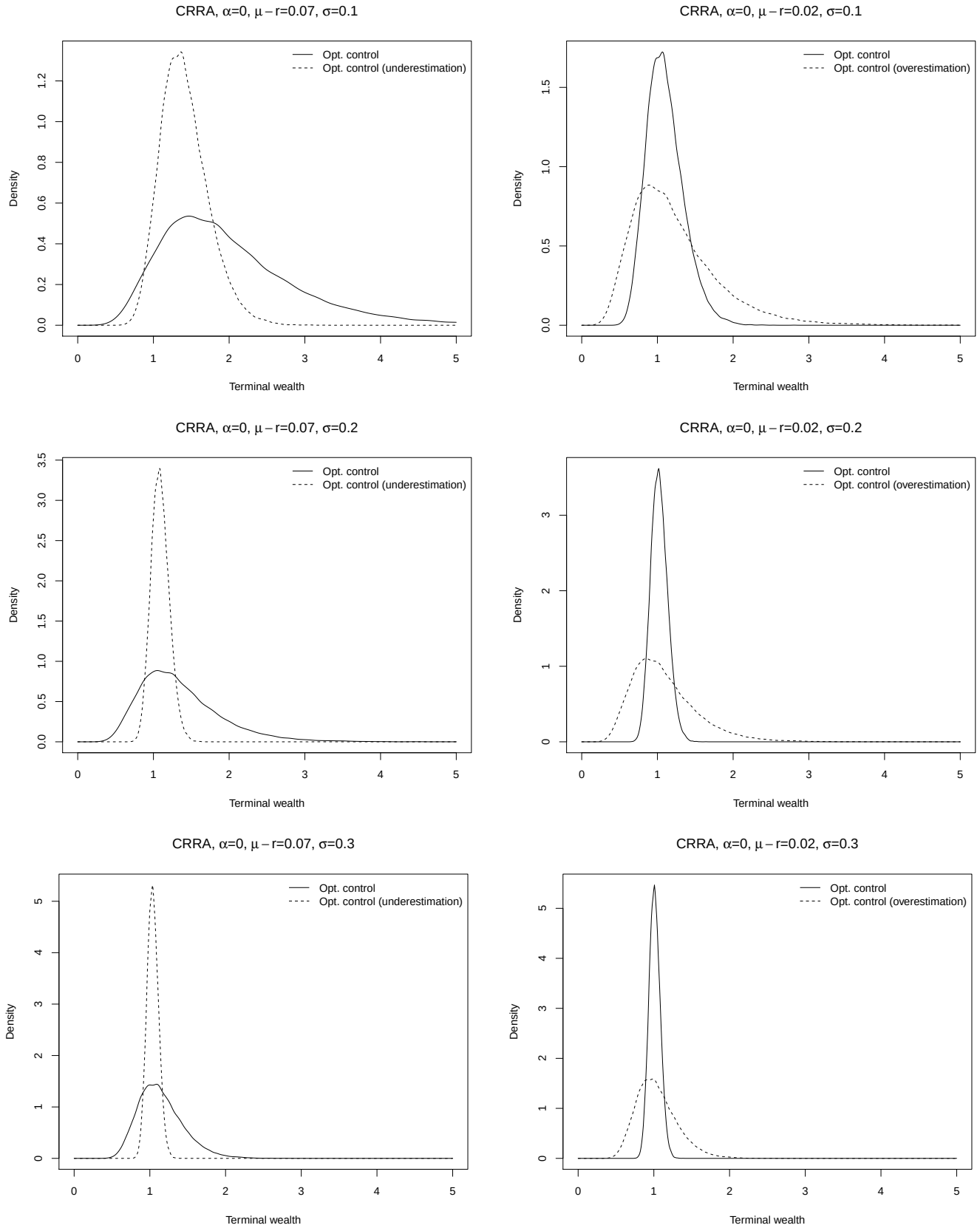

Figure 3.8: Terminal wealth distributions for the CRRA (2) investor with optimal control and optimal control calculated with mis-estimate of the risk premium, without capital protection $\alpha = 0$. The different panels correspond to different parameter sets. All amounts are simulated over 5 years and evaluated in terms of their discounted values. Distributions are based on 50,000 draws for each configuration.

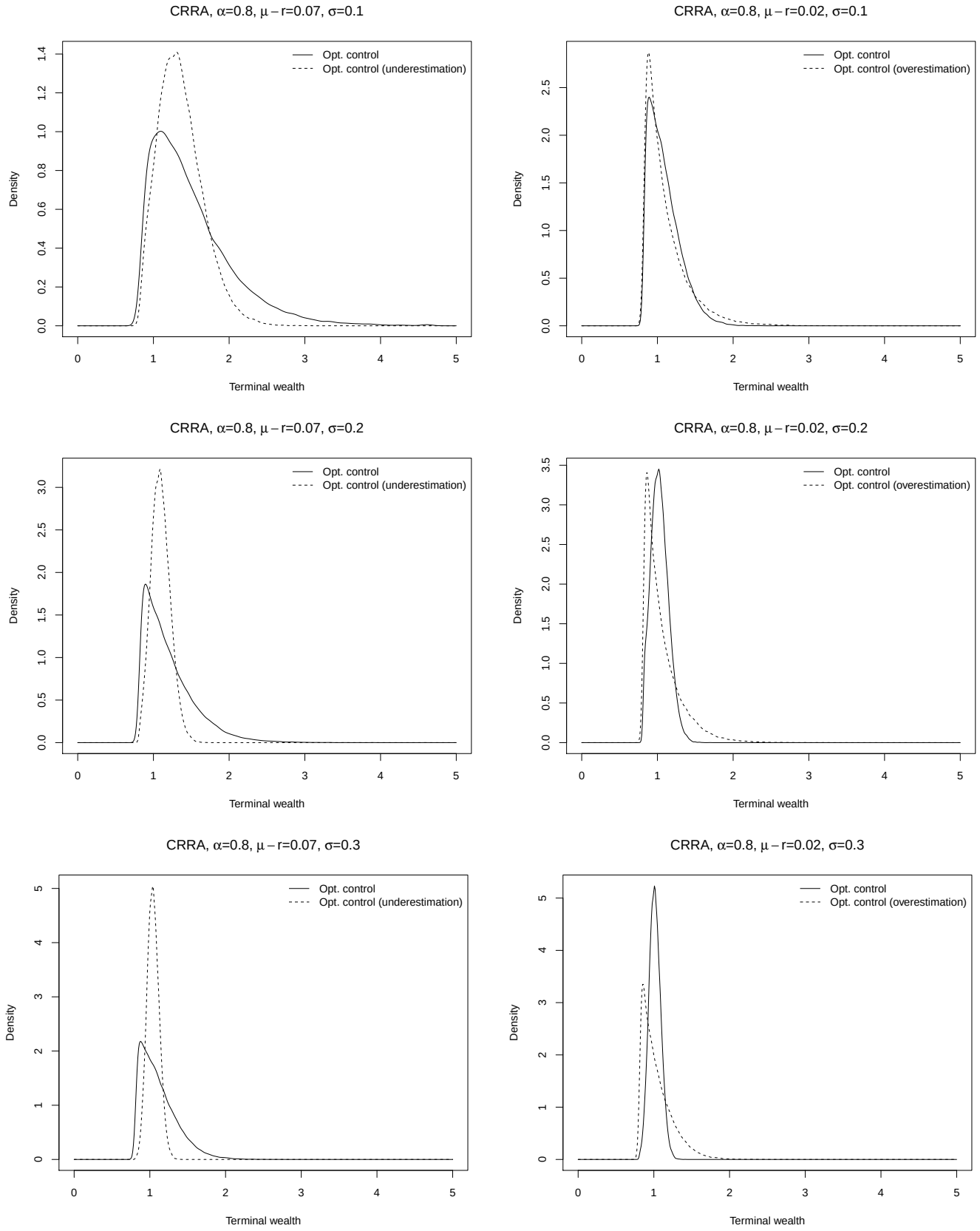

Figure 3.9: Terminal wealth distributions for the CRRA (2) investor with optimal control and optimal control calculated with mis-estimate of the risk premium, with capital protection $\alpha = 0.8$. The different panels correspond to different parameter sets. All amounts are simulated over 5 years and evaluated in terms of their discounted values. Distributions are based on 50,000 draws for each configuration.

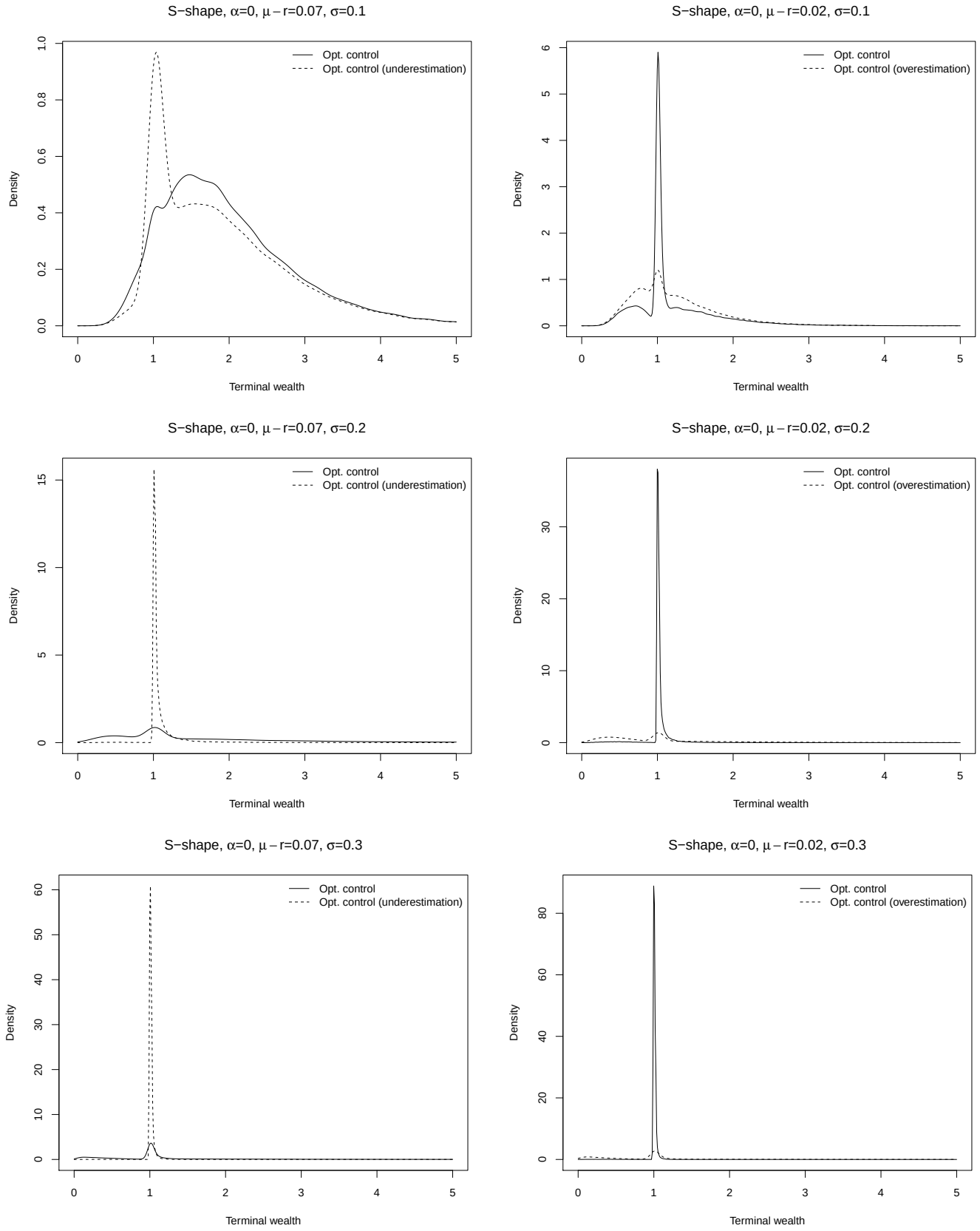

Figure 3.10: Terminal wealth distributions for the investor with S-shape utility with optimal control and optimal control calculated with mis-estimate of the risk premium, without capital protection $\alpha = 0$. The different panels correspond to different parameter sets. All amounts are simulated over 5 years and evaluated in terms of their discounted values. Distributions are based on 50,000 draws for each configuration.

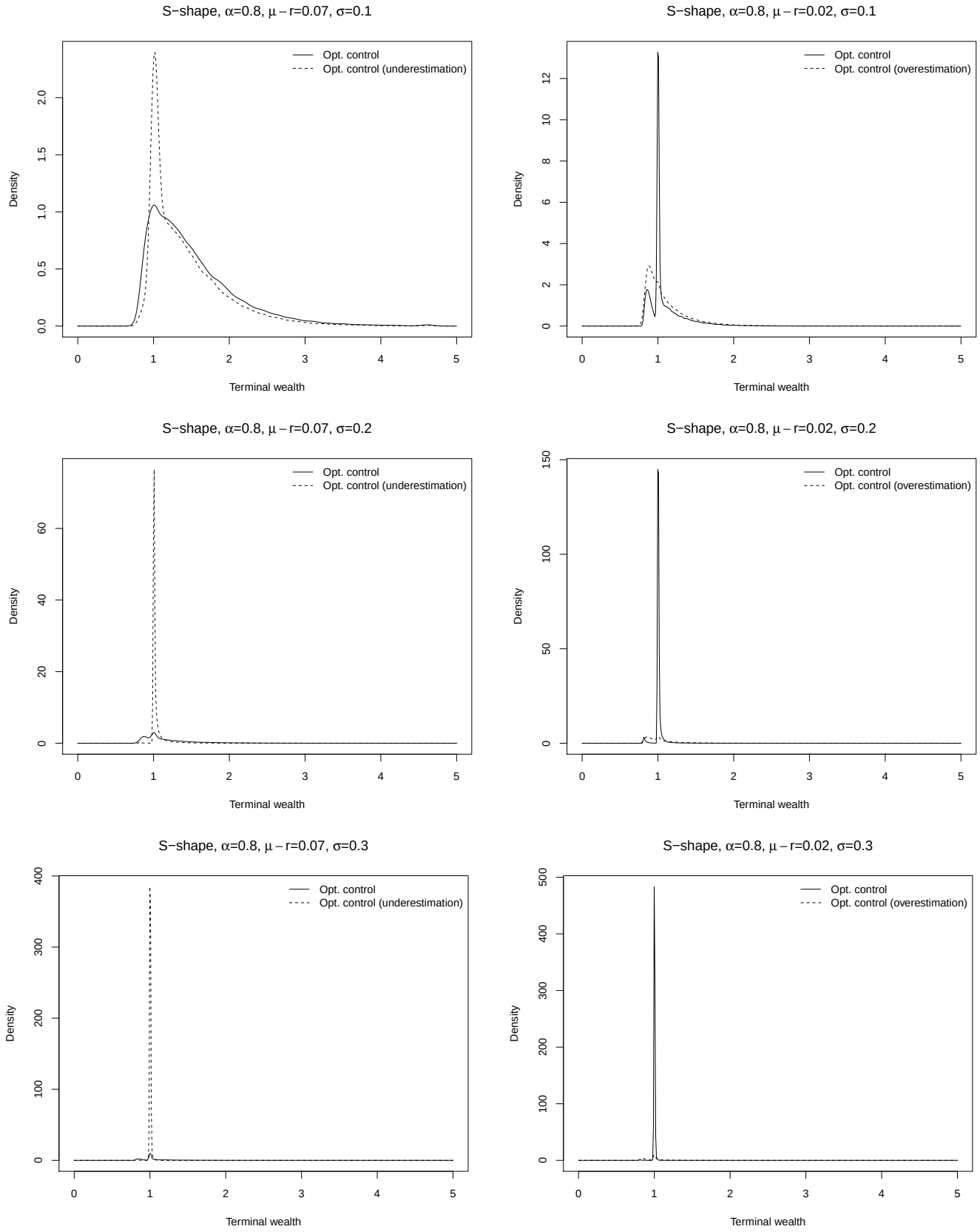

Figure 3.11: Terminal wealth distributions for the investor with S-shape utility with optimal control and optimal control calculated with mis-estimate of the risk premium, with capital protection $\alpha = 0.8$. The different panels correspond to different parameter sets. All amounts are simulated over 5 years and evaluated in terms of their discounted values. Distributions are based on 50,000 draws for each configuration.

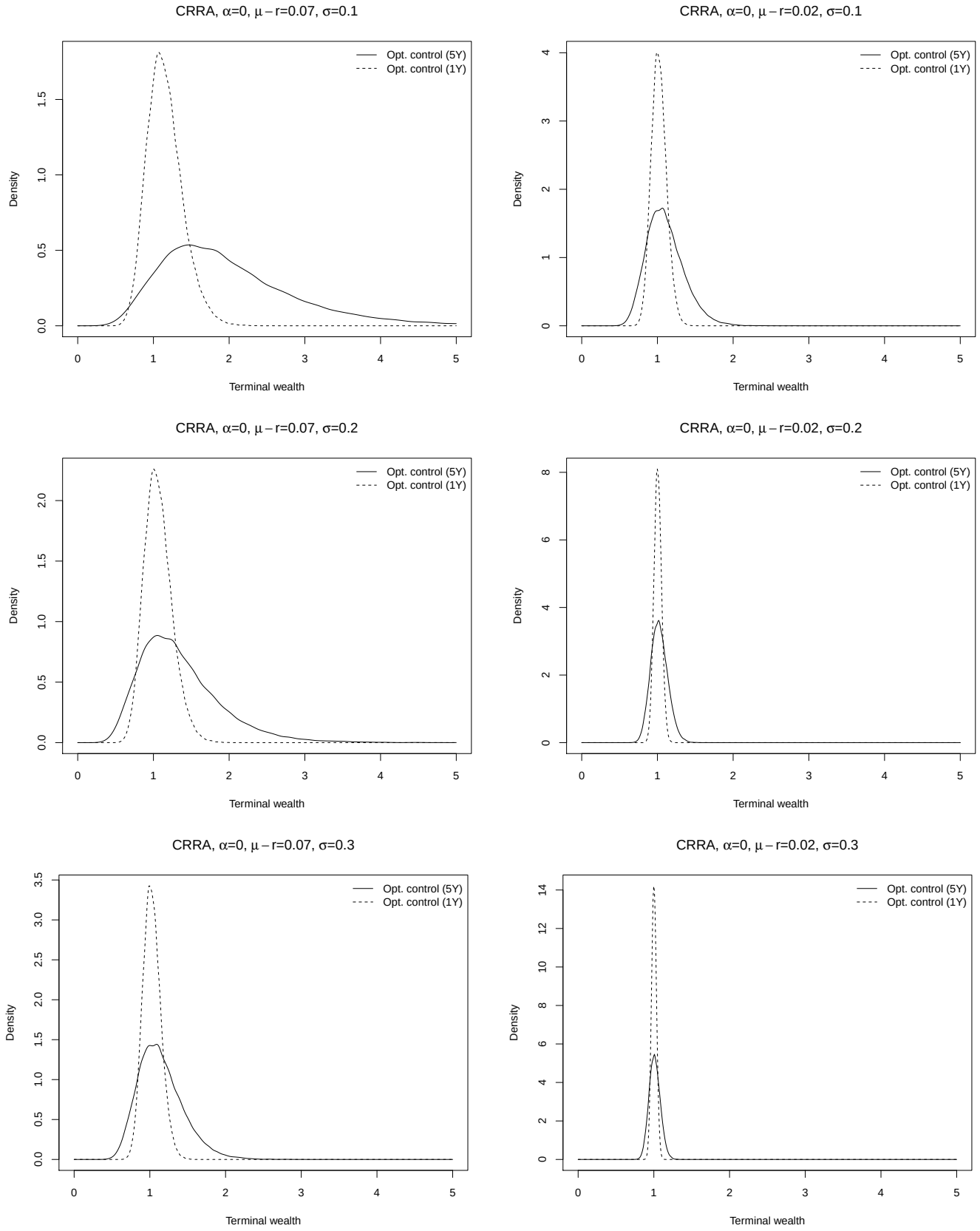

Figure 3.12: Terminal wealth distributions for the CRRA (2) investor with optimal control and without capital protection $\alpha = 0$, when the investment horizon is equal to 5 years and 1 year. The different panels correspond to different parameter sets. All amounts are simulated over the two horizons and evaluated in terms of their discounted values. Distributions are based on 50,000 draws for each configuration.

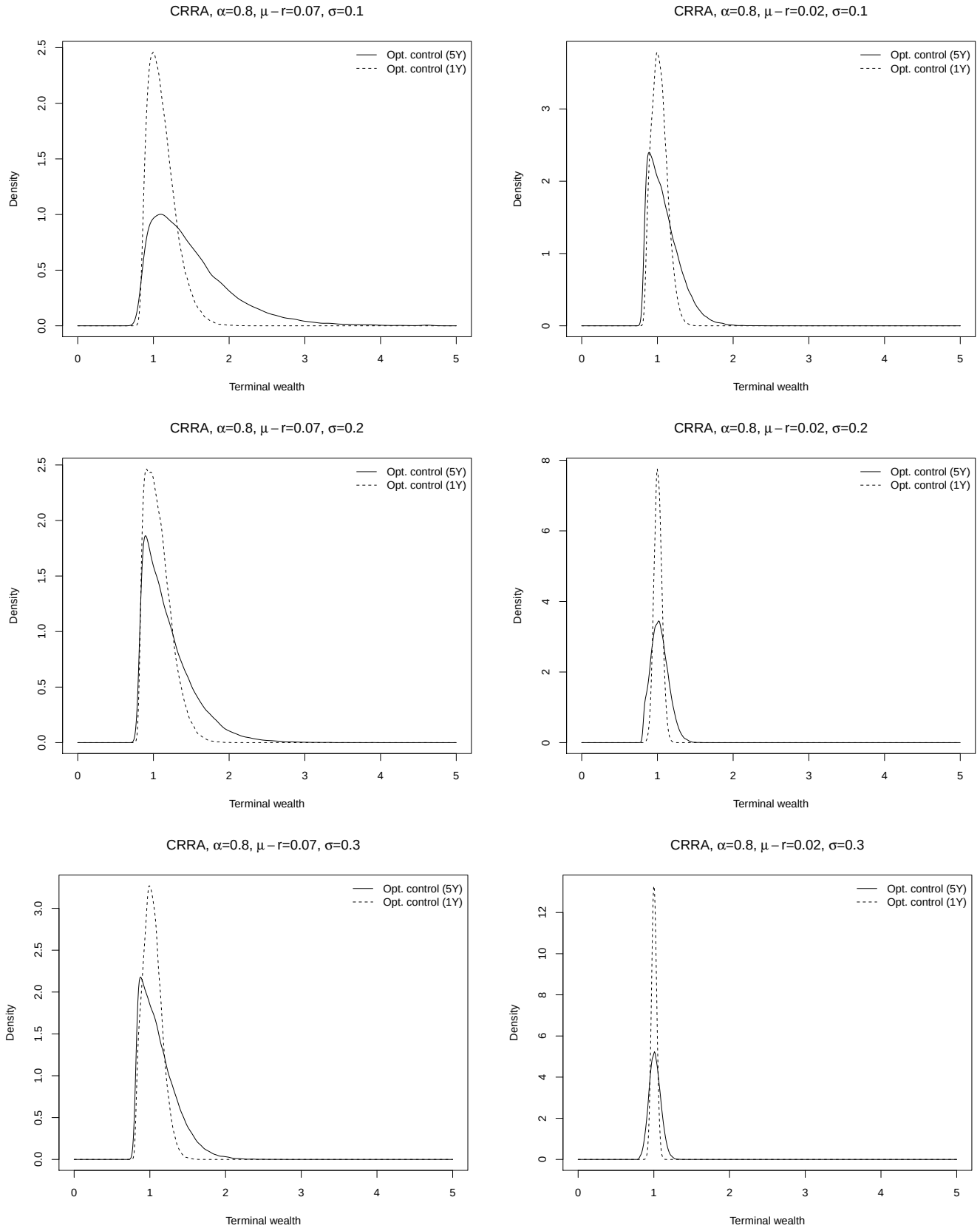

Figure 3.13: Terminal wealth distributions for the CRRA (2) investor with optimal control and with capital protection $\alpha = 0.8$, when the investment horizon is equal to 5 years and 1 year. The different panels correspond to different parameter sets. All amounts are simulated over the two horizons and evaluated in terms of their discounted values. Distributions are based on 50,000 draws for each configuration.

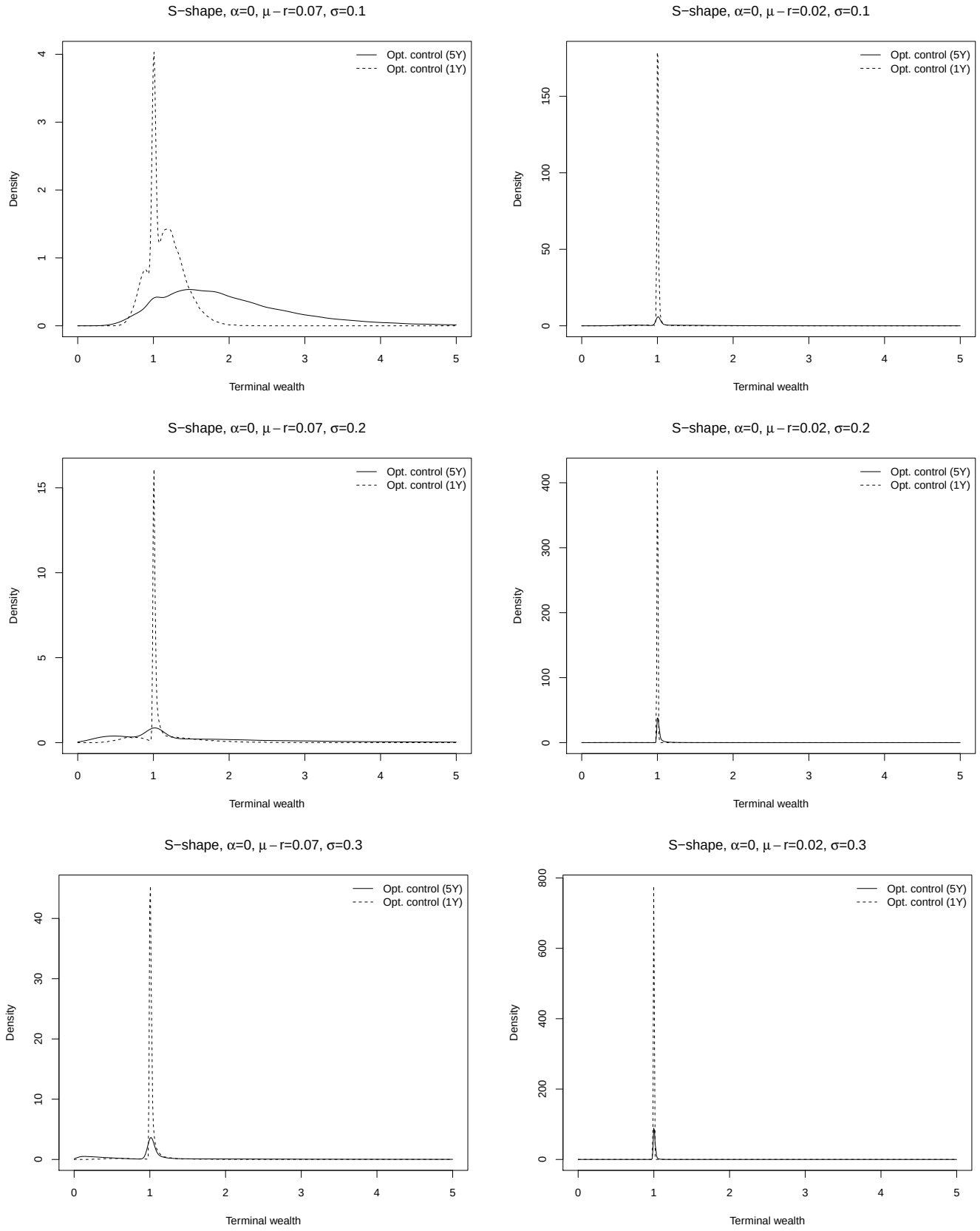

Figure 3.14: Terminal wealth distributions for the S-shape utility investor with optimal control and without capital protection $\alpha = 0$, when the investment horizon is equal to 5 years and 1 year. The different panels correspond to different parameter sets. All amounts are simulated over the two horizons and evaluated in terms of their discounted values. Distributions are based on 50,000 draws for each configuration.

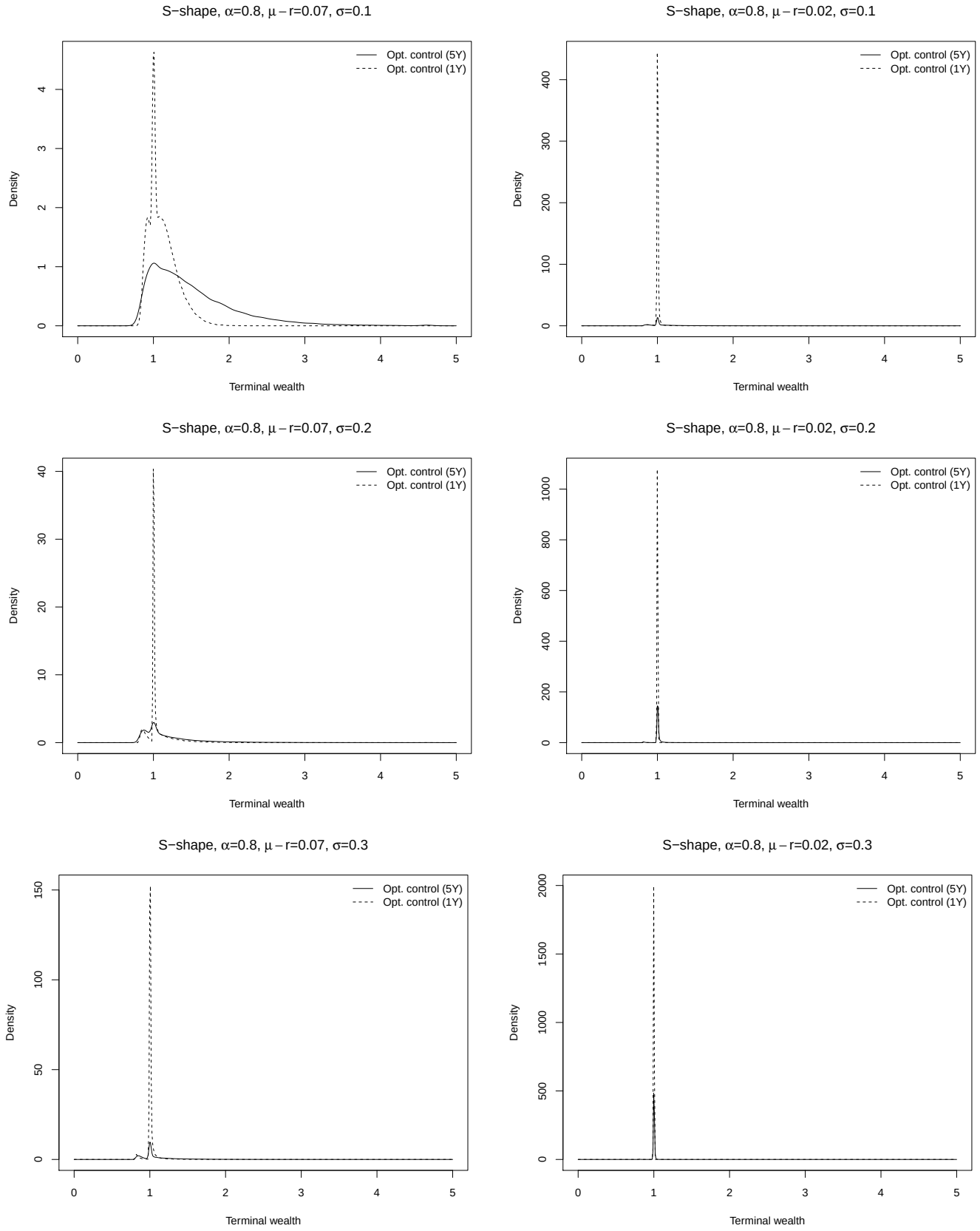

Figure 3.15: Terminal wealth distributions for the S-shape utility investor with optimal control and with capital protection $\alpha = 0.8$, when the investment horizon is equal to 5 years and 1 year. The different panels correspond to different parameter sets. All amounts are simulated over the two horizons and evaluated in terms of their discounted values. Distributions are based on 50,000 draws for each configuration.

Appendix

3.A Discretization method

In order to derive an approximation of the optimal control p^* for each point on the grid, we use the finite difference method, applied to the non linear operator $\mathcal{L}_p$. Let $v_{i,j}^k$ be a discrete approximation of $v(t_k, x_i, y_j)$ and let $\mathcal{L}_p^k$ be the discretized version of the differential operator $\mathcal{L}_p$ defined by (3.16), on the grid G. This operator can be discretized using spacial forward, centered or backward differencing in the x and y directions to give:

$$\begin{aligned}(\mathcal{L}_p^k v^k)_{i,j} = \alpha_{i,j}^k(p) v_{i-1,j}^k + \beta_{i,j}^k(p) v_{i+1,j}^k + \gamma_{i,j}^k(p) v_{i,j-1}^k + \delta_{i,j}^k(p) v_{i,j+1}^k \\ - (\alpha_{i,j}^k(p) + \beta_{i,j}^k(p) + \gamma_{i,j}^k(p) + \delta_{i,j}^k(p)) v_{i,j}^k,\end{aligned} \tag{3.29}$$

where p stands for a given value of the control and $\alpha_{i,j}^k, \beta_{i,j}^k, \gamma_{i,j}^k$ and $\delta_{i,j}^k$ are coefficient functions depending on the value of the control and the model parameters. Note that $\gamma_{i,j}^k(p) = \delta_{i,j}^k(p) = 0$ for all i, j and p, given the absence of derivatives with respect to y in (3.16). Therefore equation (3.29) reduces to:

$$(\mathcal{L}_p^k v^k)_{i,j} = \alpha_{i,j}^k(p) v_{i-1,j}^k + \beta_{i,j}^k(p) v_{i+1,j}^k - (\alpha_{i,j}^k(p) + \beta_{i,j}^k(p)) v_{i,j}^k. \tag{3.30}$$

Following Forsyth and Labahn (2008), the discretization method must maintain the monotonicity of $\mathcal{L}_p^k$ applied to $v_{i,j}^k$ in order to insure the consistency of our approach. This requirement is equivalent to

$$\alpha_{i,j}^k(p) \geq 0, \qquad \beta_{i,j}^k(p) \geq 0, \qquad \text{for all } i, j, \text{ and } p. \tag{3.31}$$

Using backward differencing for the first derivative $\frac{\left(v_{i,j}^k - v_{i-1,j}^k\right)}{\Delta x}$ and centered differencing for the second derivative $\frac{\left(v_{i+1,j}^k - 2v_{i,j}^k + v_{i-1,j}^k\right)}{\Delta x^2}$, we obtain the following coefficients:

$$\alpha_{i,j}^k(p) = -\frac{1}{\Delta x}(\mu(t) - r(t)) p (x_i - \alpha y_j), \qquad \beta_{i,j}^k(p) = \frac{1}{2\Delta x^2}\sigma(t)^2 p^2 (x_i - \alpha y_j)^2. \tag{3.32}$$

On the other hand, using forward differencing for the first derivative $\frac{\left(v_{i+1,j}^k - v_{i,j}^k\right)}{\Delta x}$ and centered differencing for the second derivative $\frac{\left(v_{i+1,j}^k - 2v_{i,j}^k + v_{i-1,j}^k\right)}{\Delta x^2}$, we obtain the following coefficients:

$$\alpha_{i,j}^k(p) = \frac{1}{2\Delta x^2}\sigma(t)^2 p^2 (x_i - \alpha y_j)^2, \qquad \beta_{i,j}^k(p) = \frac{1}{\Delta x}(\mu(t) - r(t)) p (x_i - \alpha y_j) + \alpha_{i,j}^k(p). \tag{3.33}$$

Therefore in our numerical procedure we use backward differencing for the first derivative if $(\mu(t) - r(t)) p (x_i - \alpha y_j) < 0$ and forward differencing otherwise.

3.B Wealth process simulation

From Proposition 6, the simulation of $(W_t^{x,\pi}, M_t^{x,\pi}, t \in [0,T])$ reduces to the simulation of $(w_t^{x,\pi}, m_t^{x,\pi}, t \in [0,T])$ solution of (3.5).

Assuming μ, r, σ constant on the period $[0,T]$, and assuming a portfolio rebalancing strategy π constant on $[t_k, t_{k+1}[$, with the discrete times $t_k = k\Delta t$, $k \in \mathbb{N}$, one has to simulate the random variable

$$w_t^{x,\pi} = x \exp\left[(\mu - r)\left(\Delta t \sum_{k=0}^{\eta(t)} \pi_k + (t - \eta(t))\pi_{\eta(t)}\right) - \frac{1}{2}\sigma^2\left(\Delta t \sum_{k=0}^{\eta(t)} \pi_k^2 + (t - \eta(t))\pi_{\eta(t)}^2\right)\right]$$
$$\times \exp\left[\sigma\left(\sum_{k=0}^{\eta(t)} (Z_{t_{k+1}} - Z_{t_k})\pi_k + (Z_t - Z_{\eta(t)})\pi_{\eta(t)}\right)\right]$$

where $\eta(t) = \lfloor \frac{t}{\Delta t} \rfloor \Delta t$, and

$$m_t^{x,\pi} = \sup_{0 \le s \le t} w_s^{x,\pi}.$$

An alternative iterative formulation is

$$w_{t_{k+1}}^{x,\pi} = w_{t_k}^{x,\pi} \exp\left[\Delta t\left((\mu - r)\pi_k - \frac{1}{2}\sigma^2\pi_k^2\right) + \sigma(Z_{t_{k+1}} - Z_{t_k})\pi_k\right]$$
$$m_{t_{k+1}}^{x,\pi} = \max\left(m_{t_k}^{x,\pi}, w_{t_k}^{x,\pi} \exp\left[\sup_{0 \le s \le \Delta t}\left(s\left((\mu - r)\pi_k - \frac{1}{2}\sigma^2\pi_k^2\right) + \sigma(Z_{t_k+s} - Z_{t_k})\pi_k\right)\right]\right)$$

for $k \in \mathbb{N}$. Now we use the following fact : for any reals a and c, the couple $(Z_t, \sup_{0 \le s \le t}(aZ_s + cs))$ has the same law as (U, Y) where U is Gaussian distributed, centered with variance t, and

$$Y(t,a,c) = \frac{1}{2}\left[aU + ct + a^2V + a(U + ct)^2\right]^{\frac{1}{2}}$$

with an exponential variable V with parameter $\frac{1}{2t}$ independent of U (see Lépingle, 1995).

Simulating a sequence $(w_{t_k}, Y_k = Y(\Delta t, \sigma\pi_k, (\mu - r)\pi_k - \frac{1}{2}\sigma^2\pi_k^2), k \in \mathbb{N})$, one can obtain the sequence $(m_{t_k}^{x,\pi}, k \in \mathbb{N})$ with

$$m_{t_{k+1}}^{x,\pi} = \max\left(m_{t_k}^{x,\pi}, w_{t_k}^{x,\pi} \exp(Y_k)\right).$$

3.C Results with $\alpha = 0.5$

For the sake of completeness, in this appendix we present the results obtained for strategies implementing the optimal control when the protection level α is set to 0.5, respectively for the CRRA investor (Table 3.26), and for the prospect theory investor (Table 3.27). The overall qualitative results in terms of preference and attractiveness of the drawdown constraint do not change.

CRRA, $\alpha = 0.5, \pi = p^*$	$\mu - r = 0.07$			$\mu - r = 0.02$		
	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$
Constant Equivalent	1.57	1.18	1.08	1.06	1.02	1.01
Mean terminal utility	-0.64	-0.85	-0.93	-0.94	-0.98	-0.99
Std. dev. terminal utility	0.26	0.32	0.25	0.21	0.11	0.07
Qt. 1% terminal utility	-1.38	-1.73	-1.65	-1.54	-1.27	-1.18
Qt. 50% terminal utility	-0.59	-0.79	-0.9	-0.92	-0.98	-0.99
Qt. 99% terminal utility	-0.23	-0.31	-0.48	-0.54	-0.75	-0.83
Mean Ann. return (%)	11.69	5.18	2.4	1.77	0.46	0.21
Min Ann. return (%)	-10.55	-12.77	-12.39	-11.33	-8.2	-5.56
Max Ann. return (%)	35.29	34.59	29.36	24.51	10.86	6.97
Std. dev. Ann. return	9.26	8.16	5.42	4.61	2.27	1.5
Mean Ann. volatility (%)	16.62	16.34	11.51	9.86	4.96	3.31
Min Ann. volatility (%)	7.93	7.15	7.16	6.56	4.55	3.03
Max Ann. volatility (%)	19.18	18.3	12.42	10.68	5.38	3.58
Skewness	0.1	0.09	0.06	0.05	0.03	0.02
Kurtosis	0.28	0.14	0.03	0.03	0.02	0.02
Max. drawdown (%)	24.6	29.8	23.59	20.97	11.88	8.26
Sharpe ratio	0.67	0.29	0.2	0.17	0.09	0.06
Calmar ratio	0.57	0.25	0.18	0.16	0.11	0.09
Average stock allocation (%)	163.66	81.04	38.23	98.29	24.75	11
Qt. 1% average control	3.48	1.88	0.83	2.12	0.52	0.23
Qt. 50% average control	4	2.09	0.9	2.27	0.54	0.23
Qt. 99% average control	4	3.22	1.66	3.21	0.63	0.26

Table 3.26: Performance measurements for CRRA (2) investor over 50,000 draws of portfolios. The protection level α is set to 0.5. The optimal control is calculated numerically. All amounts are simulated over 5 years and evaluated in terms of their discounted values.

S-shape, $\alpha = 0.5, \pi = p^*$	$\mu - r = 0.07$			$\mu - r = 0.02$		
	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$
Constant Equivalent	1.79	1.36	1.15	1.06	1	1
Mean terminal utility	0.81	0.4	0.19	0.08	0	0
Std. dev. terminal utility	0.75	1.05	0.88	0.51	0.12	0.02
Qt. 1% terminal utility	-0.72	-1.1	-1.14	-0.98	-0.74	-0.02
Qt. 50% terminal utility	0.72	0.17	0.02	0.02	0	0
Qt. 99% terminal utility	2.89	2.92	2.91	1.61	0.23	0.02
Mean Ann. return (%)	11.65	6.32	3.25	2	0.11	0.01
Min Ann. return (%)	-10.86	-12.72	-12.84	-11.88	-12.51	-12.68
Max Ann. return (%)	35.48	35.53	36.11	34.7	33.97	3.27
Std. dev. Ann. return	9.26	12.36	10.46	6.32	1.37	0.16
Mean Ann. volatility (%)	16.59	22.62	19.84	12.47	1.72	0.12
Min Ann. volatility (%)	7.95	6.24	0.18	0.15	0.07	0.05
Max Ann. volatility (%)	19.17	32.11	34.61	18.82	21.07	19.86
Skewness	0.1	0.16	0.16	0.09	0.24	0.09
Kurtosis	0.3	1.55	1.84	1.59	9.14	4
Max. drawdown (%)	24.58	38.25	37.2	25.89	4.13	0.31
Sharpe ratio	0.67	0.2	0.08	0.1	0.01	-0.01
Calmar ratio	0.57	0.22	0.12	0.11	0.03	0.04
Average stock allocation (%)	163.1	105.48	60.39	115.86	6.47	0.32
Qt. 1% average control	3.48	2.05	1.13	1.09	0.01	0
Qt. 50% average control	3.99	3.68	2.01	3.13	0.09	0
Qt. 99% average control	4	3.93	3.25	3.96	1.48	0.03

Table 3.27: Performance measurements for the S-shape utility investor over 50,000 draws of portfolios. The protection level α is set to 0.5. The optimal control is calculated numerically. All amounts are simulated over 5 years and evaluated in terms of their discounted values.

Bibliography

Acharya, V. V. and L. H. Pedersen (2005). Asset pricing with liquidity risk. *Journal of Financial Economics 77*(2), 375–410.

Adrian, T. and J. Rosenberg (2008). Stock returns and volatility: Pricing the short-run and long-run components of market risk. *The Journal of Finance 63*(6), 2997–3030.

Ait-Sahalia, Y. and R. Kimmel (2007). Maximum likelihood estimation of stochastic volatility models. *Journal of Financial Economics 83*(2), 413–452.

Alfonsi, A. (2005). On the discretization schemes for the cir (and bessel squared) processes. *Monte Carlo Methods and Applications 11*(4), 355.

Alizadeh, S., M. W. Brandt, and F. X. Diebold (2002). Range-based estimation of stochastic volatility models. *The Journal of Finance 57*(3), 1047–1091.

Andersen, T. G. (1996). Return volatility and trading volume: An information flow interpretation of stochastic volatility. *The Journal of Finance 51*(1), 169–204.

Andersen, T. G. and T. Bollerslev (1998). Deutsche mark–dollar volatility: intraday activity patterns, macroeconomic announcements, and longer run dependencies. *the Journal of Finance 53*(1), 219–265.

Andersen, T. G., T. Bollerslev, and F. X. Diebold (2007). Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. *The review of economics and statistics 89*(4), 701–720.

Anderson, B. D. and J. B. Moore (2012). *Optimal filtering*. Courier Dover Publications.

Anderson, T. W. (2011). *The statistical analysis of time series*, Volume 19. John Wiley & Sons.

Andreou, E. and E. Ghysels (2002). Detecting multiple breaks in financial market volatility dynamics. *Journal of Applied Econometrics 17*(5), 579–600.

Ané, T. and H. Geman (2000). Order flow, transaction clock, and normality of asset returns. *The Journal of Finance 55*(5), 2259–2284.

Ang, A. and G. Bekaert (2007). Stock return predictability: Is it there? *Review of Financial Studies 20*(3), 651.

Ang, A. and M. Piazzesi (2003). A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary economics 50*(4), 745–787.

Avramov, D., T. Chordia, and A. Goyal (2006). Liquidity and autocorrelations in individual stock returns. *The Journal of Finance 61*(5), 2365–2394.

Bachelier, L. (1900). *Théorie de la spéculation*. Gauthier-Villars.

Bajgrowicz, P., O. Scaillet, and A. Treccani (2015). Jumps in high-frequency data: Spurious detections, dynamics, and news. *Management Science*.

Baker, M. and J. Wurgler (2000). The equity share in new issues and aggregate stock returns. *the Journal of Finance 55*(5), 2219–2257.

Barberis, N. (2000). Investing for the long run when returns are predictable. *The Journal of Finance 55*(1), 225–264.

Barberis, N., M. Huang, and T. Santos (2001). Prospect theory and asset prices. *Quarterly Journal of Economics 116*(1).

Barreto, H. and F. Howland (2006). *Introductory econometrics: using Monte Carlo simulation with Microsoft excel*. Cambridge University Press.

Barrieu, P. and N. El Karoui (2009). Pricing, hedging, and designing derivatives with risk measures., indifference pricing: Theory and applications.

Basak, S. (1995). A general equilibrium model of portfolio insurance. *Review of Financial studies 8*(4), 1059–1090.

Basak, S. and A. Shapiro (2001). Value-at-risk-based risk management: optimal policies and asset prices. *Review*

of Financial studies 14(2), 371–405.

Benartzi, S. and R. H. Thaler (1995). Myopic loss aversion and the equity premium puzzle. *The Quarterly Journal of Economics*.

Berkaoui, A., M. Bossy, A. Diop, et al. (2008). Euler scheme for sdes with non-lipschitz diffusion coefficient: strong convergence. *ESAIM: Probability and Statistics 12*.

Berkelaar, A. B., R. Kouwenberg, and T. Post (2004). Optimal portfolio choice under loss aversion. *Review of Economics and Statistics 86*(4), 973–987.

Bernard, C. and M. Ghossoub (2010). Static portfolio choice under cumulative prospect theory. *Mathematics and financial economics 2*(4), 277–306.

Bessembinder, H. and P. J. Seguin (1993). Price volatility, trading volume, and market depth: Evidence from futures markets. *Journal of financial and Quantitative Analysis 28*(01), 21–39.

Białkowski, J., S. Darolles, and G. Le Fol (2008). Improving vwap strategies: A dynamic volume approach. *Journal of Banking & Finance 32*(9), 1709–1722.

Binsbergen, V., H. Jules, and R. S. Koijen (2010). Predictive regressions: A present-value approach. *The Journal of Finance 65*(4), 1439–1471.

Black, F. and A. Perold (1992). Theory of constant proportion portfolio insurance. *Journal of Economic Dynamics and Control 16*(3-4), 403–426.

Black, F. and R. Rouhani (1989). Constant proportion portfolio insurance and the synthetic put option: a comparison. *Institutional Investor focus on Investment Management, edited by Frank J. Fabozzi. Cambridge, Mass.: Ballinger*, 695–708.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics 31*(3), 307–327.

Bollerslev, T., L. Xu, and H. Zhou (2015). Stock return and cash flow predictability: The role of volatility risk. *Journal of Econometrics 187*(2), 458–471.

Bossy, M., A. Diop, et al. (2007). An efficient discretisation scheme for one dimensional sdes with a diffusion coefficient function of the form| x|ˆ a, a in [1/2, 1).

Bouchard, B., R. Elie, and C. Imbert (2010). Optimal control under stochastic target constraints. *SIAM Journal on Control and Optimization 48*(5), 3501–3531.

Bouchard, B., R. Elie, and N. Touzi (2009). Stochastic target problems with controlled loss. *SIAM Journal on Control and Optimization 48*(5), 3123–3150.

Boyle, P. and W. Tian (2007). Portfolio management with constraints. *Mathematical Finance 17*(3), 319–343.

Brandt, M. W. and Q. Kang (2004). On the relationship between the conditional mean and volatility of stock returns: A latent var approach. *Journal of Financial Economics 72*(2), 217–257.

Brennan, M. J. and E. S. Schwartz (1989). Portfolio insurance and financial market equilibrium. *Journal of Business*, 455–472.

Brennan, M. J., R. Solanki, et al. (1981). Optimal portfolio insurance. *Journal of Financial and Quantitative Analysis 16*(3), 279–300.

Brigo, D. and A. Alfonsi (2005). Credit default swap calibration and derivatives pricing with the ssrd stochastic intensity model. *Finance and Stochastics 9*(1), 29–42.

Campbell, J., S. Grossman, and J. Wang (1993). Trading volume and serial correlation in stock returns. *Quarterly Journal of Economics 108*(4), 905–939.

Campbell, J. and R. Shiller (1988). The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies 1*(3), 195–228.

Campbell, J. Y. (1987). Stock returns and the term structure. *Journal of Financial Economics 18*(2), 373–399.

Campbell, J. Y. (1991). A variance decomposition for stock returns. Technical report, National Bureau of Economic Research.

Campbell, J. Y. and J. Ammer (1993). What moves the stock and bond markets? a variance decomposition for long-term asset returns. *The Journal of Finance 48*(1), 3–37.

Campbell, J. Y. and S. B. Thompson (2008). Predicting excess stock returns out of sample: Can anything beat the historical average?. *Review of Financial Studies 21*(4), 1509 – 1531.

Carassus, L., M. Rásonyi, and A. M. Rodrigues (2015). Non-concave utility maximisation on the positive real axis in discrete time. *Mathematics and Financial Economics 9*(4), 325–349.

Carter, C. K. and R. Kohn (1994). On gibbs sampling for state space models. *Biometrika 81*(3), 541–553.

Carvalho, C. M., H. F. Lopes, and R. E. McCulloch (2015). On the long run volatility of stocks.

Chernov, M., A. R. Gallant, E. Ghysels, and G. Tauchen (2003). Alternative models for stock price dynamics. *Journal of Econometrics 116*(1), 225–257.

Cherny, V. and J. Obłój (2013). Portfolio optimisation under non-linear drawdown constraints in a semimartingale financial model. *Finance and Stochastics 17*(4), 771–800.

Clark, P. K. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica 41*(1), 135–155.

Cochrane, J. (2008). The dog that did not bark: A defense of return predictability. *Review of Financial Studies 21*(4), 1533.

Cochrane, J. H. (2005). Time series for macroeconomics and finance. *Manuscript, University of Chicago*.

Colacito, R., R. F. Engle, and E. Ghysels (2011). A component model for dynamic correlations. *Journal of Econometrics 164*(1), 45–59.

Conrad, J. and G. Kaul (1988). Time-variation in expected returns. *Journal of Business*, 409–425.

Cont, R. and P. Tankov (2009). Constant proportion portfolio insurance in the presence of jumps in asset prices. *Mathematical Finance 19*(3), 379–401.

Cooper, M. (1999). Filter rules based on price and volume in individual security overreaction. *Review of Financial Studies 12*(4), 901–935.

Cox, J. and H. Leland (1982). Notes on intertemporal investment policies. *Unpublished working paper*.

Cox, J. C., J. E. Ingersoll Jr, and S. A. Ross (1985). A theory of the term structure of interest rates. *Econometrica: Journal of the Econometric Society*, 385–407.

Cujean, J. and M. Hasler (2014). Why does return predictability concentrate in bad times?

Cvitanic, J. and I. Karatzas (1995). On portfolio optimization under" drawdown" constraints. *IMA Volumes in Mathematics and its Applications 65*, 35–35.

Da, Z., Q. Liu, and E. Schaumburg (2013). A closer look at the short-term return reversal. *Management Science 60*(3), 658–674.

Dangl, T. and M. Halling (2012). Predictive regressions with time-varying coefficients. *Journal of Financial Economics*.

Darolles, S., G. Le Fol, and G. Mero (2015). Measuring the liquidity part of volume. *Journal of Banking & Finance 50*, 92–105.

De Brouwer, P. and F. Van den Spiegel (2001). The fallacy of large numbers revisited: The construction of a utility function that leads to the acceptance of two games, while one is rejected. *Journal of Asset Management 1*(3), 257–266.

Dichtl, H. and W. Drobetz (2011). Portfolio insurance and prospect theory investors: Popularity and optimal design of capital protected financial products. *Journal of Banking & Finance 35*(7), 1683–1697.

Dierkes, M., C. Erner, and S. Zeisberger (2010). Investment horizon and the attractiveness of investment strategies: A behavioral approach. *Journal of Banking & Finance 34*(5), 1032–1046.

Ding, Z. and C. W. Granger (1996). Modeling volatility persistence of speculative returns: a new approach. *Journal of econometrics 73*(1), 185–215.

Duffee, G. R. (2007). Are variations in term premia related to the macroeconomy. *Unpublished working paper. University of California, Berkeley*.

Duffie, D. (2010). *Dynamic asset pricing theory*. Princeton University Press.

Ebert, S., B. Koos, and J. C. Schneider (2012). On the optimal type and level of guarantees for prospect theory investors. In *Paris December 2012 Finance Meeting EUROFIDAI-AFFI Paper*.

El Karoui, N., M. Jeanblanc, and V. Lacoste (2005). Optimal portfolio management with american capital guarantee. *Journal of Economic Dynamics and Control 29*(3), 449–468.

El Karoui, N. and A. Meziou (2006). Constrained optimization with respect to stochastic dominance: application to portfolio insurance. *Mathematical Finance 16*(1), 103–117.

Elie, R. (2006). *Contrôle stochastique et méthodes numériques en finance mathématique*. Ph. D. thesis, ENSAE ParisTech.

Elie, R. (2008). Finite time merton strategy under drawdown constraint: a viscosity solution approach. *Applied Mathematics and Optimization 58*(3), 411–431.

Elie, R. and N. Touzi (2008). Optimal lifetime consumption and investment under a drawdown constraint. *Finance and Stochastics 12*(3), 299–330.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica 50*(4), 987–1008.

Engle, R. F. (1990). Stock volatility and the crash of'87: Discussion. *Review of Financial Studies 3*(1), 103–06.

Engle, R. F., E. Ghysels, and B. Sohn (2013). Stock market volatility and macroeconomic fundamentals. *Review of Economics and Statistics 95*(3), 776–797.

Engle, R. F. and C. W. Granger (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica: journal of the Econometric Society*, 251–276.

Engle, R. F. and G. Lee (1999). A permanent and transitory component model of stock return volatility. *Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive WJ Granger*, 475–497.

Engle, R. F. and J. G. Rangel (2008). The spline-garch model for low-frequency volatility and its global macroeconomic causes. *Review of Financial Studies 21*(3), 1187–1222.

Engle, R. F. and M. E. Sokalska (2012). Forecasting intraday volatility in the us equity market. multiplicative component garch. *Journal of Financial Econometrics 10*(1), 54–83.

Epps, T. W. and M. L. Epps (1976). The stochastic dependence of security price changes and transaction volumes: Implications for the mixture-of-distributions hypothesis. *Econometrica 44*(2), 305–21.

Estep, T. and M. Kritzman (1988). Tipp. *The Journal of Portfolio Management 14*(4), 38–42.

Fama, E. F. and K. R. French (1988). Dividend yields and expected stock returns. *Journal of Financial Economics 22*(1), 3–25.

Fama, E. F. and K. R. French (1989). Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics 25*(1), 23–49.

Feller, W. (1951). Two singular diffusion problems. *Annals of mathematics*, 173–182.

Ferson, W. E. and C. R. Harvey (1991). The variation of economic risk premiums. *Journal of Political Economy*, 385–415.

Fleming, J., C. Kirby, and B. Ostdiek (2005). Arch effects and trading volume. *Working paper*.

Fleming, J., C. Kirby, and B. Ostdiek (2006). Stochastic volatility, trading volume, and the daily flow of information. *The Journal of Business 79*(3), 1551–1590.

Forsyth, P. and G. Labahn (2008). Numerical methods for controlled hamilton-jacobi-bellman pdes in finance. *Journal of Computational Finance 11*(2), 1–44.

Foster, F. D. and S. Viswanathan (1995). Can speculative trading explain the volume–volatility relation? *Journal of Business & Economic Statistics 13*(4), 379–396.

Frühwirth-Schnatter, S. (1994). Data augmentation and dynamic linear models. *Journal of time series analysis 15*(2), 183–202.

Gallant, A. R., C.-T. Hsu, and G. Tauchen (1999). Using daily range data to calibrate volatility diffusions and extract the forward integrated variance. *Review of Economics and Statistics 81*(4), 617–631.

Gallant, A. R., P. E. Rossi, and G. Tauchen (1992). Stock prices and volume. *Review of Financial studies 5*(2), 199–242.

Ghysels, E., A. C. Harvey, and E. Renault (1996). 5 stochastic volatility. *Handbook of statistics 14*, 119–191.

Glosten, L. R., R. Jagannathan, and D. E. Runkle (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance 48*(5), 1779–1801.

Gneiting, T. (2011). Making and evaluating point forecasts. *Journal of the American Statistical Association 106*(494), 746–762.

Goetzmann, W. N. and P. Jorion (1993). Testing the predictive power of dividend yields. *The Journal of Finance 48*(2), 663–679.

Gomes, F. J. (2005). Portfolio choice and trading volume with loss-averse investors. *The Journal of Business 78*(2), 675–706.

Gouriéroux, C. and J. Jasiak (2006). Autoregressive gamma processes. *Journal of Forecasting 25*(2), 129–152.

Goyal, A. and I. Welch (2003). Predicting the equity premium with dividend ratios. *Management Science 49*(5), 639–654.

Goyal, A. and I. Welch (2008). A Comprehensive Look at The Empirical Performance of Equity Premium Prediction. *Review of Financial Studies 21*(4), 1455.

Graham, B., D. L. F. Dodd, and S. Cottle (1934). *Security analysis*. McGraw-Hill New York.

Grinblatt, M. and B. Han (2005). Prospect theory, mental accounting, and momentum. *Journal of financial economics 78*(2), 311–339.

Grossman, S. J. and J.-L. Vila (1989). Portfolio insurance in complete markets: A note. *Journal of Business*, 473–476.

Grossman, S. J. and Z. Zhou (1993). Optimal investment strategies for controlling drawdowns. *Mathematical*

Finance 3(3), 241–276.

Grossman, S. J. and Z. Zhou (1996). Equilibrium analysis of portfolio insurance. *The Journal of Finance 51*(4), 1379–1403.

Grüne, L. and W. Semmler (2008). Asset pricing with loss aversion. *Journal of Economic Dynamics and Control 32*(10), 3253–3274.

Haigh, M. S. and J. A. List (2005). Do professional traders exhibit myopic loss aversion? an experimental analysis. *The Journal of Finance 60*(1), 523–534.

Hamidi, B., B. Maillet, and J.-L. Prigent (2014). A dynamic autoregressive expectile for time-invariant portfolio protection strategies. *Journal of Economic Dynamics and Control 46*, 1–29.

Hamilton, J. D. and G. Lin (1996). Stock market volatility and the business cycle. *Journal of Applied Econometrics 11*(5), 573–593.

Hansen, B. E. (1996). Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica: Journal of the econometric society*, 413–430.

Harris, M. and A. Raviv (1993). Differences of opinion make a horse race. *Review of Financial studies 6*(3), 473–506.

He, X. D. and X. Y. Zhou (2011). Portfolio choice under cumulative prospect theory: An analytical treatment. *Management Science 57*(2), 315–331.

Heckman, J. J. (1979). *Statistical models for discrete panel data*. Department of Economics and Graduate School of Business, University of Chicago.

Henkel, S. J., J. S. Martin, and F. Nardari (2011). Time-varying short-horizon predictability. *Journal of Financial Economics 99*(3), 560–580.

Hiemstra, C. and J. D. Jones (1994). Testing for linear and nonlinear granger causality in the stock price-volume relation. *The Journal of Finance 49*(5), 1639–1664.

Hsieh, D. A. (1989). Modeling heteroscedasticity in daily foreign-exchange rates. *Journal of Business & Economic Statistics 7*(3), 307–317.

Jin, H. and X. Yu Zhou (2008). Behavioral portfolio selection in continuous time. *Mathematical Finance 18*(3), 385–426.

Johannes, M., A. Korteweg, and N. Polson (2014). Sequential learning, predictability, and optimal portfolio returns. *The Journal of Finance 69*(2), 611–644.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, 263–291.

Karpoff, J. M. (1987). The relation between price changes and trading volume: A survey. *Journal of Financial and quantitative Analysis 22*(01), 109–126.

Keim, D. B. and R. F. Stambaugh (1986). Predicting returns in the stock and bond markets. *Journal of Financial Economics 17*(2), 357–390.

Kelly, B. and S. Pruitt (2013). Market expectations in the cross-section of present values. *The Journal of Finance 68*(5), 1721–1756.

Khuman, A., N. Constantinou, and S. Phelps (2012). Constant proportion portfolio insurance strategies under cumulative prospect theory with reference point adaptation. *Available at SSRN 2109551*.

Kim, C.-J. and C. R. Nelson (1999). State-space models with regime switching: classical and gibbs-sampling approaches with applications. *MIT Press Books 1*.

Klass, M. J. and K. Nowicki (2005). The grossman and zhou investment strategy is not always optimal. *Statistics & probability letters 74*(3), 245–252.

Kvašňáková, K. (2013). Is imperfection better? evidence from predicting stock and bond returns. *Unpublished working paper. Vienna University of Economics and Business*.

Kwiatkowski, D., P. C. Phillips, P. Schmidt, and Y. Shin (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of econometrics 54*(1), 159–178.

Lamoureux, C. G. and W. D. Lastrapes (1990a). Heteroskedasticity in stock return data: volume versus garch effects. *The Journal of Finance 45*(1), 221–229.

Lamoureux, C. G. and W. D. Lastrapes (1990b). Persistence in variance, structural change, and the garch model. *Journal of Business & Economic Statistics 8*(2), 225–234.

Lamoureux, C. G. and W. D. Lastrapes (1994). Endogenous trading volume and momentum in stock-return volatility. *Journal of Business & Economic Statistics 12*(2), 253–260.

Lamoureux, C. G. and G. Zhou (1996). Temporary components of stock returns: What do the data tell us? *Review of Financial Studies 9*(4), 1033–1059.

Lattimore, P. K., J. R. Baker, and A. D. Witte (1992). The influence of probability on risky choice: A parametric examination. *Journal of Economic Behavior & Organization 17*(3), 377–400.

Leland, H. and M. Rubinstein (1976). Portfolio insurance: A guide to dynamic hedging. *The evolution of portfolio insurance*.

Leland, H. E. (1980). Who should buy portfolio insurance? *The Journal of Finance 35*(2), 581–594.

Lépingle, D. (1995). Euler scheme for reflected stochastic differential equations. *Mathematics and Computers in Simulation 38*(1), 119–126.

Lettau, M. and S. Ludvigson (2001). Consumption, aggregate wealth, and expected stock returns. *Journal of Finance*, 815–849.

Lettau, M. and S. van Nieuwerburgh (2008). Reconciling the return predictability evidence. *Review of Financial Studies 21*(4), 1607.

Liesenfeld, R. (1998). Dynamic bivarsate mixture models: Modeling the behavior of prices and trading volume. *Journal of Business & Economic Statistics 16*(1), 101–109.

Llorente, G., R. Michaely, G. Saar, and J. Wang (2002). Dynamic volume-return relation of individual stocks. *Review of Financial studies 15*(4), 1005–1047.

Lo, A. W. and J. Wang (2000). Trading volume: definitions, data analysis, and implications of portfolio theory. *Review of Financial Studies 13*(2), 257–300.

Maheu, J. M. and T. H. McCurdy (2004). News arrival, jump dynamics, and volatility components for individual stock returns. *The Journal of Finance 59*(2), 755–793.

Mantilla-Garcia, D. and V. Vaidyanathan (2011). Predicting Stock Returns in the presence of Uncertain Structural Breaks and Sample Noise. *Working Paper, Edhec Business School*.

McCracken, M. (2007). Asymptotics for out of sample tests of Granger causality. *Journal of Econometrics 140*(2), 719–752.

Merton, R. (1995). Influence of mathematical models in finance on practice: past, present, and future. *Mathematical Models in Finance*, 1–14.

Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *The review of Economics and Statistics 51*(3), 247–257.

Merton, R. C. (1980). On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics 8*(4), 323–361.

Merton, R. C. (1993). On the microeconomic theory of investment under uncertainty. *Handbook of Mathematical Economics 2*, 601–669.

Meucci, A. (2011). "P" versus "Q": Differences and commonalities between the two areas of quantitative finance. *The Quant Classroom by Attilio Meucci - GARP Risk Professional February 2011*, 47–50.

Murphy, K. M. and R. H. Topel (2002). Estimation and inference in two-step econometric models. *Journal of Business & Economic Statistics 20*(1), 88–97.

Nelson, D. B. (1990). Arch models as diffusion approximations. *Journal of econometrics 45*(1), 7–38.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, 347–370.

Pages, G., F. Panloup, et al. (2009). Approximation of the distribution of a stationary markov process with application to option pricing. *Bernoulli 15*(1), 146.

Pastor, L. and R. F. Stambaugh (2003). Liquidity risk and expected stock returns. *Journal of Political Economy 111*(3).

Pástor, L. and R. F. Stambaugh (2009). Predictive systems: Living with imperfect predictors. *The Journal of Finance 64*(4), 1583–1628.

Pástor, L. and R. F. Stambaugh (2012). Are stocks really less volatile in the long run? *The Journal of Finance 67*(2), 431–478.

Pesaran, M. H. and A. Timmermann (1995). Predictability of stock returns: Robustness and economic significance. *The Journal of Finance 50*(4), 1201–1228.

Pettenuzzo, D. and F. Ravazzolo (2014). Optimal portfolio choice under decision-based model combinations. *Available at SSRN*.

Pettenuzzo, D., A. Timmermann, and R. Valkanov (2014). Forecasting stock returns under economic constraints. *Journal of Financial Economics (forthcoming)*.

Piatti, I. and F. Trojani (2014). Dividend growth predictability and the price-dividend ratio. *Available at SSRN*.

Piatti, I. and F. Trojani (2015). Predictable risks and predictive regression in present-value models. *Available at SSRN 1786897*.

Rapach, D., J. Strauss, and G. Zhou (2010). Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *Review of Financial Studies 23*(2), 821.

Richardson, M. and T. Smith (1994). A direct test of the mixture of distributions hypothesis: Measuring the daily flow of information. *Journal of Financial and Quantitative Analysis 29*(01), 101–116.

Ross, S. A. (1999). Adding risks: Samuelson's fallacy of large numbers revisited. *Journal of Financial and Quantitative Analysis 34*(03), 323–339.

Rousseeuw, P. J. and K. V. Driessen (1999). A fast algorithm for the minimum covariance determinant estimator. *Technometrics 41*(3), 212–223.

Rytchkov, O. (2012). Filtering out expected dividends and expected returns. *The Quarterly Journal of Finance 2*(03).

Samuelson, P. (1963). Risk and uncertainty: A fallacy of large numbers. *Scientia 98*(4-5), 108–113.

Schwert, G. W. (1989). Why does stock market volatility change over time? *The Journal of Finance 44*(5), 1115–1153.

Shalen, C. T. (1993). Volume, volatility, and the dispersion of beliefs. *Review of Financial Studies 6*(2), 405–434.

Stambaugh, R. F. (1986). Bias in Regressions with Lagged Stochastic Regressors. *Working Paper University of Chicago*.

Stambaugh, R. F. (1999). Predictive Regressions. *Journal of Financial Economics 54*, 375–ñ421.

Tauchen, G. E. and M. Pitts (1983). The price variability-volume relationship on speculative markets. *Econometrica 51*(2), 485–505.

Tversky, A. and D. Kahneman (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty 5*(4), 297–323.

Van Binsbergen, J. H. and R. S. Koijen (2011). Likelihood-based estimation of exactly-solved present-value models. *Available at SSRN 1439849*.

Vrecko, D. and N. Branger (2009). Why is portfolio insurance attractive to investors? *Available at SSRN 1519344*.

Wang, J. (1994). A model of competitive stock trading volume. *Journal of Political Economy 102*(1), 127–68.

Watanabe, T. (2000). Bayesian analysis of dynamic bivariate mixture models: Can they explain the behavior of returns and trading volume? *Journal of Business & Economic Statistics 18*(2), 199–210.

Yang, M. (2008). Normal log-normal mixture, leptokurtosis and skewness. *Applied Economics Letters 15*(9), 737–742.

Ying, C. C. (1966). Stock market prices and volumes of sales. *Econometrica 34*(3), 676–685.

Résumé

Cette thèse présente trois contributions indépendantes. La première partie se concentre sur la modélisation de la moyenne conditionnelle des rendements du marché actions : le rendement espéré du marché. Ce dernier est souvent modélisé à l'aide d'un processus AR(1). Cependant, des études montrent que lors de mauvaises périodes économiques la prédictibilité des rendements est plus élevée. Etant donné que le modèle AR(1) exclut par construction cette propriété, nous proposons d'utiliser un modèle CIR. Les implications sont étudiées dans le cadre d'un modèle espace-état bayésien. La deuxième partie est dédiée à la modélisation de la volatilité des actions et des volumes de transaction. La relation entre ces deux quantités a été justifiée par l'hypothèse de mélange de distribution (MDH). Cependant, cette dernière ne capture pas la persistance de la variance, à la différence des spécifications GARCH. Nous proposons un modèle à deux facteurs combinant les deux approches, afin de dissocier les variations de volatilité court terme et long terme. Le modèle révèle plusieurs régularités importantes sur la relation volume-volatilité. La troisième partie s'intéresse à l'analyse des stratégies d'investissement optimales sous contrainte "drawdown". Le problème étudié est celui de la maximisation d'utilité à horizon fini pour différentes fonctions d'utilité. Nous calculons les stratégies optimales en résolvant numériquement l'équation de Hamilton-Jacobi-Bellman, qui caractérise le principe de programmation dynamique correspondant. En se basant sur un large panel d'expérimentations numériques, nous analysons les divergences des allocations optimales.

Mots clés : marché financier, modèle espace-état, filtre de Kalman, analyse bayésienne, volatilité stochastique, modèle GARCH, optimisation de portefeuille, contrôle stochastique, finance comportementale.

Abstract

This PhD thesis presents three independent contributions. The first part is concentrated on the modeling of the conditional mean of stock market returns: the expected market return. The latter is often modeled as an AR(1) process. However, empirical studies have found that during bad times return predictability is higher. Given that the AR(1) model excludes by construction this property, we propose to use instead a CIR model. The implications of this specification are studied within a flexible Bayesian state-space model. The second part is dedicated to the modeling of stocks volatility and trading volume. The empirical relationship between these two quantities has been justified by the Mixture of Distribution Hypothesis (MDH). However, this framework notably fails to capture the obvious persistence in stock variance, unlike GARCH specifications. We propose a two-factor model of volatility combining both approaches, in order to disentangle short-run from long-run volatility variations. The model reveals several important regularities on the volume-volatility relationship. The third part of the thesis is concerned with the analysis of optimal investment strategies under the drawdown constraint. The finite horizon expectation maximization problem is studied for different types of utility functions. We compute the optimal investments strategies, by solving numerically the Hamilton–Jacobi–Bellman equation, that characterizes the dynamic programming principle related to the stochastic control problem. Based on a large panel of numerical experiments, we analyze the divergences of optimal allocation programs.

Keywords : stock market, state-space model, Kalman filter, Bayesian analysis, stochastic volatility, GARCH model, portfolio optimization, stochastic control, behavioral finance.