The role of fat-tails, multiple variance components, and pricing kernels in option pricing

by

Kadir Gökhan Babaoğlu

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Graduate Department of Rotman School of Management
University of Toronto

© Copyright 2016 by Kadir Gökhan Babaoğlu
Abstract

The role of fat-tails,
multiple variance components,
and pricing kernels in option pricing

Kadir Gökhan Babaoğlu
Doctor of Philosophy
Graduate Department of Rotman School of Management
University of Toronto
2016

My dissertation, composed of two chapters, explores the pricing of index and individual equity options contracts. These chapters make three modeling choices on (i) state variables, (ii) return innovations and (iii) the pricing kernel, and answer the question about what we can learn from stocks and options data.

Both chapters specify a variance-dependent pricing kernel, which allows non-monotonicity when projected onto returns. While first chapter employs Inverse Gaussian distribution to capture fat-tailed dynamics of returns, second chapter chooses to model distribution of returns as a normal shock plus Compound Poisson jumps. Regarding the state variables, chapter 1 uses long-run and short-run variance components, whereas chapter 2 defines normal and jump variance components as the state variables.

The first chapter nests multiple volatility components, fat tails and a variance-dependent pricing kernel in a single option model and compare their contribution to describing returns and option data. All three features lead to statistically significant model improvements. A variance-dependent pricing kernel is economically most important and improves option fit by 17% on average and more so for two-factor models. A second volatility component improves the option fit by 9% on average. Fat tails improve option fit by just over 4% on average, but more so when a variance-dependent pricing kernel is applied. Overall these three model features are complements rather than substitutes: the importance of one feature increases in conjunction with the others.

Focusing on individual equity options, second chapter develops a new factor model that explores (i) if a separate beta for market jumps is needed, (ii) cross-sectional differences in jump betas of stocks, and (iii) the role of jump betas in explaining equity option prices. Differentiating between normal beta and jump beta, the model predicts that a stock with higher sensitivity to market jumps (normal shocks) have higher out-of-the-money (at-the-money) option prices. The results show that jump betas are needed to adequately explain equity options.
I am indebted to Peter Christoffersen for his guidance. For helpful comments I would like to thank Jason Wei and Kris Jacobs. I also thank Fousseni Chabi-Yo, Jin-Chuan Duan, Mathieu Fournier, Andras Fulop, Michael Hasler, Tom McCurdy, Chayawat Ornthalailai. Financial support from the University of Toronto is gratefully acknowledged.
Contents

Abstract ii

List of Tables v

List of Figures vi

List of Appendices viii

1 Option Valuation with Volatility Components, Fat Tails, and Non-Monotonic Pricing Kernels 1

1.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1

1.2 A Class of GARCH Dynamics for Option Valuation . . . . . . . . . . . . . . . . . . . . . . . . . 3

1.2.1 The GARCH(1,1) Model . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 4

1.2.2 The GARCH(2,2) Model and a Component Model . . . . . . . . . . . . . . . . . . . . . . 4

1.2.3 The IG-GARCH(1,1) Model . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6

1.2.4 The IG-GARCH Component Model . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 7

1.2.5 The Gaussian Limit of the IG Model . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 9

1.3 The Risk-Neutral Model and Option Valuation . . . . . . . . . . . . . . . . . . . . . . . . . . . . 10

1.3.1 Risk-Neutralization . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 10

1.3.2 Preference Parameters and Risk-Neutral Parameters . . . . . . . . . . . . . . . . . . . . . 11

1.3.3 Nested Option Models . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 12

1.3.4 Option Valuation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 13

1.4 Data and Estimation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 14

1.4.1 Data . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 14

1.4.2 Estimation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 15

1.5 Empirical Results . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 17

1.5.1 Fitting Returns and Fitting Options . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 18

1.5.2 Sequential Estimation of the Non-Monotonic Pricing Kernel Parameter . . . . . . . . . . . 19

1.5.3 Capturing Dynamics in Higher Moments . . . . . . . . . . . . . . . . . . . . . . . . . . 20

1.5.4 The Relative Importance of Model Features for Option RMSE . . . . . . . . . . . . . . . 22

1.5.5 Capturing Smiles and Smirks . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 23

1.5.6 Model-Implied Relative Risk Aversion . . . . . . . . . . . . . . . . . . . . . . . . . . . . 24

1.6 Conclusion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 26

1.7 Figures . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 27

1.8 Tables . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 34
## The pricing of market jumps in the cross-section of equity options

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Introduction</td>
<td>48</td>
</tr>
<tr>
<td>2.2 Data and preliminary analysis</td>
<td>50</td>
</tr>
<tr>
<td>2.2.1 Description of data</td>
<td>50</td>
</tr>
<tr>
<td>2.2.2 Is jump beta equal to normal beta?</td>
<td>51</td>
</tr>
<tr>
<td>2.3 Model</td>
<td>54</td>
</tr>
<tr>
<td>2.3.1 Dynamics of variance components</td>
<td>54</td>
</tr>
<tr>
<td>2.3.2 The pricing kernel</td>
<td>56</td>
</tr>
<tr>
<td>2.3.3 The risk premium</td>
<td>56</td>
</tr>
<tr>
<td>2.3.4 Option pricing</td>
<td>58</td>
</tr>
<tr>
<td>2.4 Estimating the betas</td>
<td>59</td>
</tr>
<tr>
<td>2.4.1 Estimating the model on market returns</td>
<td>59</td>
</tr>
<tr>
<td>2.4.2 Maximum likelihood estimation of stock return dynamics</td>
<td>60</td>
</tr>
<tr>
<td>2.4.3 Challenges in the estimations</td>
<td>60</td>
</tr>
<tr>
<td>2.4.4 Betas and option fit</td>
<td>61</td>
</tr>
<tr>
<td>2.5 Results</td>
<td>61</td>
</tr>
<tr>
<td>2.5.1 Market return dynamics</td>
<td>62</td>
</tr>
<tr>
<td>2.5.2 Equity return dynamics</td>
<td>63</td>
</tr>
<tr>
<td>2.5.3 Evidence from options</td>
<td>64</td>
</tr>
<tr>
<td>2.6 Conclusion</td>
<td>65</td>
</tr>
<tr>
<td>2.7 Figures</td>
<td>66</td>
</tr>
<tr>
<td>2.8 Tables</td>
<td>71</td>
</tr>
</tbody>
</table>

**Bibliography**

92
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Returns and Options Data</td>
<td>35</td>
</tr>
<tr>
<td>1.2</td>
<td>Maximum Likelihood Estimation Results for Return Distribution</td>
<td>36</td>
</tr>
<tr>
<td>1.3</td>
<td>Maximum Likelihood Estimation Results for Risk Neutral Distribution</td>
<td>37</td>
</tr>
<tr>
<td>1.4</td>
<td>Sequential Maximum Likelihood Estimation</td>
<td>38</td>
</tr>
<tr>
<td>1.5</td>
<td>Implied Volatility RMSE and Bias by Moneyness</td>
<td>39</td>
</tr>
<tr>
<td>1.6</td>
<td>Implied Volatility RMSE and Bias by Maturity</td>
<td>40</td>
</tr>
<tr>
<td>2.1</td>
<td>Descriptive statistics</td>
<td>72</td>
</tr>
<tr>
<td>2.2</td>
<td>Regression results</td>
<td>73</td>
</tr>
<tr>
<td>2.3</td>
<td>Estimation results of market return and volatility dynamics</td>
<td>74</td>
</tr>
<tr>
<td>2.4</td>
<td>Estimation results of equity return and volatility dynamics</td>
<td>75</td>
</tr>
<tr>
<td>2.5</td>
<td>Estimation results of equity options</td>
<td>76</td>
</tr>
</tbody>
</table>
## List of Figures

1.1 RMSE and Option Likelihood Values versus $\xi$ ................................................. 27  
1.2 Spot Variance Paths Using Return-Based Estimates ............................................. 28  
1.3 Leverage Correlation and Volatility of Variance Using Return-Based Estimates ........ 29  
1.4 Term Structure of Variance, Skewness and Kurtosis ............................................. 30  
1.5 Spot Variance Paths Using Option-Based Estimates ............................................. 31  
1.6 Leverage Correlation and Volatility of Variance Using Option-Based Estimates ....... 32  
1.7 Model-Based Implied Volatility Smiles in IG-GARCH Component Model ................. 33  
2.1 Standardized CAPM-betas and variance ratios. ...................................................... 66  
2.2 $\beta_{y,j}$, $\beta_{y,j}$ vs. $\mu_{z,j}$, $\mu_{y,j}$ ................................................................. 67  
2.3 Jump and diffusive $\beta$ vs. implied volatilities. ................................................... 68  
2.4 Filtered S&P 500 index return shocks ............................................................... 69  
2.5 Filtered state variables: Volatility of normal shocks and jumps of S&P 500 index returns ............................... 70
### List of Appendices

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appendix 1.A</td>
<td>Martingale Restrictions</td>
<td>41</td>
</tr>
<tr>
<td>Appendix 1.B</td>
<td>The Risk-Neutral Distribution</td>
<td>43</td>
</tr>
<tr>
<td>Appendix 1.C</td>
<td>The Risk-Neutral Component Model</td>
<td>45</td>
</tr>
<tr>
<td>Appendix 1.D</td>
<td>Conditional Moments</td>
<td>46</td>
</tr>
<tr>
<td>Appendix 2.A</td>
<td>Model</td>
<td>77</td>
</tr>
<tr>
<td>Appendix 2.B</td>
<td>Model details</td>
<td>79</td>
</tr>
<tr>
<td>Appendix 2.C</td>
<td>Pricing kernel</td>
<td>80</td>
</tr>
<tr>
<td>Appendix 2.D</td>
<td>Solution</td>
<td>81</td>
</tr>
<tr>
<td>Appendix 2.E</td>
<td>Risk-neutralization</td>
<td>84</td>
</tr>
<tr>
<td>Appendix 2.F</td>
<td>Moment-generating function of market returns</td>
<td>86</td>
</tr>
<tr>
<td>Appendix 2.G</td>
<td>Moment-generating function of stock returns</td>
<td>88</td>
</tr>
<tr>
<td>Appendix 2.H</td>
<td>Option prices</td>
<td>91</td>
</tr>
</tbody>
</table>
Chapter 1

Option Valuation with Volatility Components, Fat Tails, and Non-Monotonic Pricing Kernels

with Peter Christoffersen, Steven Heston and Kris Jacobs

1.1 Introduction

By accounting for heteroskedasticity and volatility clustering, empirical studies on option valuation substantially improve on the Black-Scholes (1973) model prices through the parametric modeling of stochastic volatility (SV), see for example Heston (1993) and Bakshi, Cao, and Chen (1997). The literature has focused on two improvements to capture the stylized facts in the data. First, by accounting for more than one volatility component, the model becomes more flexible and its modeling of the term structure of volatility improves. This approach is advocated by Duffie, Pan, and Singleton (2000) and implemented on option prices by, among others, Bates (2000), Christoffersen, Heston, and Jacobs (2009), and Xu and Taylor (1994).\footnote{See e.g. Chernov, Gallant, Ghysels, and Tauchen (2003) for a study of multiple volatility components in the underlying return series, and Ronnibouts and Stentoft (2014 and 2015) for normal mixture models.} Christoffersen, Jacobs, Ornthanalai, and Wang (2008) propose a discrete-time GARCH option valuation model with two volatility components which has more structure, by modeling total volatility as evolving around a stochastic long-run mean.
The second modeling improvement that reliably improves model fit is to augment stochastic volatility with jumps in returns and/or volatility. A large number of studies have implemented this approach. Intuitively, the advantage offered by jump processes is that they allow for conditional nonnormality, and therefore for instantaneous skewness and kurtosis. In discrete-time modeling, an equivalent approach uses innovations that are conditionally non-Gaussian. Examples of this approach are Christoffersen, Heston, and Jacobs (2006), who use Inverse Gaussian innovations, and Barone-Adesi, Engle, and Mancini (2008) who take a nonparametric approach.

The studies cited above demonstrate convincingly that these two modeling approaches improve model fit for both the option prices and the underlying returns. However, the most important challenge faced by these models is the simultaneous modeling of the underlying returns and the options. This position is forcefully articulated by for example Bates (1996b, 2003). Stentoft (2008), and Andersen, Fusari, and Todorov (2015) address this by fitting realized (physical) volatility together with option prices, but this still leaves open the question of a pricing kernel that links the observed “physical” measure and the risk-neutral measure inherent in option prices. In particular, deficiencies in a model’s ability to simultaneously describe returns and option prices may not exclusively be due to the specification of the driving process, but could also be caused by a misspecified price of risk, or equivalently the pricing kernel.

The literature focuses on pricing kernels that depend on wealth, originating in the seminal work of Brennan (1979) and Rubinstein (1976). Liu, Pan, and Wang (2004) discuss the specification of the price of risk when SV models are augmented with Poisson jumps. Several papers, including Ait-Sahalia and Lo (1998), Jackwerth (2000), Rosenberg and Engle (2002), Bakshi, Madan, and Panayotov (2010), Brown and Jackwerth (2012), and Chabi-Yo (2012), have documented deviations from and explored extensions to the traditional log-linear pricing kernel. In recent work, Christoffersen, Heston, and Jacobs (2013) specify a more general pricing kernel that depends on volatility as well as wealth. The kernel is non-monotonic after projecting onto wealth, which is consistent with recent evidence by Cuesdeanu and Jackwerth (2015). Christoffersen, Heston, and Jacobs (2013) show that the more general pricing kernel provides a superior fit to option prices and returns.

The literature thus suggests at least three important improvements on the benchmark SV option pricing model. First, multiple volatility components; second, conditional nonnormality or jumps; and third, non-monotonic pricing kernels. It is important to nest these features within a common framework in order to have a “horserace” comparison of their importance. In addition, examining these features jointly shows how they interact in describing returns and options. Ideally these different model features ought to be complements rather than substitutes. The

---


3 Linn, Shive, and Shumway (2014) argue that the finding of a nonmonotone pricing kernel could be an artifact of the econometric method used. Cuesdeanu and Jackwerth (2015) show that the finding of a nonmonotone kernel is robust across a range of econometric techniques.
second volatility factor should improve the modeling of the volatility term structure, and therefore the valuation of options of different maturities, and long-maturity options in particular. Non-Gaussian innovations should prove most useful to capture the moneyness dimension for short-maturity options, which is usually referred to as the smirk. The non-monotonic (variance-dependent) pricing kernel has an entirely different purpose, because its relevance lies in the joint modeling of index returns and options, rather than the modeling of options alone.

However, the existing literature does not contain any evidence on whether these model features are indeed complements when confronted with the data. The literature does also not address the question of which model feature is statistically and economically most significant. This paper is the first to address this issue by comparing the three features within a nested model. We conduct an extensive empirical evaluation of the three model features using returns data, using options data, and finally using a sequential estimation exercise. We find that all three model features lead to statistically significant model improvements. A variance-dependent pricing kernel is economically most important and improves option fit by 17% on average and more so for two-factor models. A second volatility factor improves the option fit by 9% on average. Fat tails improve option fit by just over 4% on average, but more so when a variance-dependent pricing kernel is applied. Our results suggest that the three features are complements rather than substitutes.

The paper proceeds as follows. Section 2 introduces a class of GARCH dynamics for option valuation. The most general return dynamic we consider has non-normal innovations and two variance components, one of which is a stochastic long-run mean. We also derive the Gaussian limit of this return process. Section 3 discusses the risk-neutralization of this process. Section 4 discusses data and estimation, and Section 5 presents the empirical results. Section 6 concludes.

1.2 A Class of GARCH Dynamics for Option Valuation

This section introduces a general class of GARCH dynamics for index option valuation. The most general model we consider is a two-factor fat-tail GARCH model with dynamics that may appear non-standard. We therefore first introduce two better-known models, and subsequently introduce the more general GARCH(2,2) model with Inverse Gaussian innovations, IG-GARCH(2,2). We also show how the IG-GARCH(2,2) model can be transformed into a component model and demonstrate how it nests the simpler cases.

One can use observable state variables to value options in any dynamic model. For example, one might use implied volatilities extracted from option prices. Alternatively, one might use a filtering technique such as the particle filter, or rely on realized volatility computed from intraday returns as in Andersen, Fusari, and Todorov (2015).
We choose a GARCH model because we want to assess whether option prices are consistent with observed returns. In this framework filtering is straightforward, which facilitates investigating the relationship between option prices and return dynamics. The GARCH approach allows us to impose economic restrictions based on observed returns, without an auxiliary filter that is separate from the assumptions of the option model. The limitation of a GARCH approach is that it does not allow one-step-ahead volatility to evolve independently of returns. This is not a significant problem in practice, because the model allows innovations in variance to be imperfectly correlated with daily (or higher frequency) returns.

The IG-GARCH processes has the continuous-time limits of the standard stochastic volatility model of Heston (1993) or pure jump process with stochastic intensity at the same time. It allows for conditional skewness in addition to heteroskedasticity of returns. Analytical tractability of Inverse Gaussian distribution helps estimation of the model on option contracts.4

1.2.1 The GARCH(1,1) Model

The first model we consider is the Heston-Nandi (2000) Gaussian GARCH(1,1) process:

\[
\begin{align*}
\ln(S(t + \Delta)) &= \ln(S(t)) + r + \mu h(t + \Delta) + \sqrt{h(t + \Delta)} z(t + \Delta), \\
 h(t + \Delta) &= \omega + \beta_1 h(t) + \alpha_1 (z(t) - \gamma_1 \sqrt{h(t)})^2.
\end{align*}
\]

(1.1a)  
(1.1b)

where \(z(t)\) has a standard normal distribution. This model allows for quasi-closed form valuation of European options, and has therefore been estimated and tested in several empirical applications.5

1.2.2 The GARCH(2,2) Model and a Component Model

A straightforward generalization of the Heston-Nandi (2000) GARCH(1,1) dynamic in (1.1a)-(1.1b) is the following GARCH(2,2) process with normal innovations:

\[
\begin{align*}
\ln(S(t + \Delta)) &= \ln(S(t)) + r + \mu h(t + \Delta) + \sqrt{h(t + \Delta)} z(t + \Delta), \\
 h(t + \Delta) &= \omega + \beta_1 h(t) + \beta_2 h(t - \Delta) \\
&\quad + \alpha_1 (z(t) - \gamma_1 \sqrt{h(t)})^2 + \alpha_2 (z(t - \Delta) - \gamma_2 \sqrt{h(t - \Delta)})^2,
\end{align*}
\]

4For a detailed discussion, see Christoffersen, Heston, and Jacobs (2006)
5See for example, Hsieh and Ritchken (2005), Barone-Adesi, Engle and Mancini (2008), and Christoffersen, Jacobs, Ornthanalai and Wang (2008).
The GARCH(2,2) model is not typically used in empirical work. However, building on Engle and Lee (1999), by imposing some parameter restrictions it can be written as the component model of Christoffersen et al. (2008):

\begin{align}
h(t + \Delta) &= q(t + \Delta) + \rho_1 (h(t) - q(t)) + \nu_h(t), \quad (1.3a) \\
q(t + \Delta) &= \omega_q + \rho_2 q(t) + \nu_q(t). \quad (1.3b)
\end{align}

where

\[
\nu_i(t) = \alpha_i [(z(t) - \gamma_i \sqrt{h(t)})^2 - 1 - \gamma_i^2 h(t)] \quad i = h, q,
\]

\[
\gamma_h = -\frac{\rho_1 \gamma_1 + \alpha_2 \gamma_2}{(\rho_2 - \rho_1) \alpha_h}, \quad \gamma_q = \frac{\rho_2 \gamma_1 + \alpha_2 \gamma_2}{(\rho_2 - \rho_1) \alpha_q},
\]

\[
\alpha_h = -\frac{\rho_1}{\rho_2 - \rho_1} \alpha_1 - \frac{1}{\rho_2 - \rho_1} \alpha_2, \quad \alpha_q = \frac{-\rho_2}{\rho_2 - \rho_1} \alpha_1 + \frac{1}{\rho_2 - \rho_1} \alpha_2,
\]

and \(\rho_1\) and \(\rho_2\) are the respective smaller and larger roots of the quadratic equation

\[
\rho^2 - (\beta_1 + \alpha_1 \gamma_1^2) \rho - \beta_2 - \alpha_2 \gamma_2^2 = 0,
\]

The dynamic for the long-run component can equivalently be expressed as

\[
q(t + \Delta) = \sigma^2 + \rho_2 (q(t) - \sigma^2) + \nu_q(t),
\]

where \(\sigma^2\) is the unconditional variance.

The component parameters can also be inverted to recover the GARCH(2,2) parameters

\[
\alpha_1 = \alpha_h + \alpha_q, \quad \alpha_2 = -\rho_2 \alpha_h - \rho_1 \alpha_q, \\
\beta_1 = \rho_1 + \rho_2 - \alpha_1 \gamma_1^2, \quad \beta_2 = -\rho_1 \rho_2 - \alpha_2 \gamma_2^2.
\]

The component structure helps interpreting the model. The coefficients of the lagged variables (in the long- and short-run components) are the roots of the process’ characteristic equation. These parameters are more informative about the process than the parameters in the GARCH(2,2) model, which facilitates estimation including the identification of appropriate parameter starting values.
1.2.3 The IG-GARCH(1,1) Model

Another generalization of the Heston-Nandi (2000) GARCH(1,1) dynamic in (1.1a)-(1.1b) is the IG-GARCH(1,1) process in Christoffersen, Heston, and Jacobs (2006), given by:

\[
\begin{align*}
\ln(S(t+\Delta)) &= \ln(S(t)) + r + \mu h(t + \Delta) + \eta y(t + \Delta), \\
h(t + \Delta) &= w + b_1 h(t) + c_1 y(t) + a_1 h(t)^2/y(t),
\end{align*}
\]

(1.4a)

(1.4b)

where \(y(t + \Delta)\) has an Inverse Gaussian distribution with degrees of freedom \(h(t + \Delta)/\eta^2\). Note that while \(y(t + \Delta)\) is a positive random variable, returns are shifted by \(\mu h(t + \Delta)\) and can have both negative and positive values. The Inverse Gaussian innovation and its reciprocal have the following conditional means

\[
\begin{align*}
E_t[y(t + \Delta)] &= h(t + \Delta)/\eta^2, \\
E_t[1/y(t + \Delta)] &= \eta^2/h(t + \Delta) + \eta^4/(h(t + \Delta))^2.
\end{align*}
\]

(1.5a)

(1.5b)

The dynamic (1.4a)-(1.4b) can be written in terms of zero-mean innovations as follows

\[
\begin{align*}
\ln(S(t+\Delta)) &= \ln(S(t)) + r + \tilde{\mu} h(t + \Delta) + \sqrt{h(t + \Delta)} z(t + \Delta), \\
h(t + \Delta) &= \tilde{w} + \tilde{b}_1 h(t) + \nu_1(t),
\end{align*}
\]

(1.6a)

(1.6b)

where

\[
\begin{align*}
\tilde{\mu} &= \mu + \eta^{-1}, \\
\tilde{w} &= w + a_1 \eta^4, \\
\tilde{b}_1 &= b_1 + c_1/\eta^2 + a_1 \eta^2, \\
z(t) &= \frac{\eta y(t) - h(t)/\eta}{\sqrt{h(t)}} \\
\nu_1(t) &= c_1 y(t) + a_1 h(t)^2/y(t) - c_1 h(t)/\eta^2 - a_1 \eta^2 h(t) - a_1 \eta^4.
\end{align*}
\]

(1.7a)

(1.7b)

(1.7c)

(1.7d)

(1.7e)

The conditional means of return and variance are given by

\[
\begin{align*}
E_t[\ln(S(t+\Delta)/S(t))] &= r + \tilde{\mu} h(t + \Delta), \\
E_t[h(t + 2\Delta)] &= \tilde{w} + \tilde{b}_1 h(t + \Delta).
\end{align*}
\]

(1.8a)

(1.8b)
The advantage of the IG-GARCH(1,1) process in (1.4a)-(1.4b) over the GARCH(1,1) process in (1.1a)-(1.1b) is that the innovation is non-normal, thus allowing for conditional skewness and kurtosis. Because the Inverse Gaussian distribution converges to the normal distribution as degree of freedom tends to infinity\(^6\), the Heston-Nandi (2000) dynamic is nested by the specification of Christoffersen, Heston, and Jacobs (2006).

### 1.2.4 The IG-GARCH Component Model

We now combine the two generalizations of the Heston-Nandi (2000) model, the Inverse Gaussian innovations and the component structure. Consider the IG-GARCH(2,2) process given by:

\[
\begin{align*}
\ln(S(t+\Delta)) &= \ln(S(t)) + r + \mu h(t + \Delta) + \eta y(t + \Delta), \quad (1.9a) \\
h(t + \Delta) &= w + b_1 h(t) + b_2 h(t - \Delta) + c_1 y(t) + c_2 y(t - \Delta) \\
&\quad + a_1 h(t)^2/y(t) + a_2 h(t - \Delta)^2/y(t - \Delta), \quad (1.9b)
\end{align*}
\]

The dynamic (1.9a)-(1.9b) can be written in terms of zero-mean innovations as follows

\[
\begin{align*}
\ln(S(t+\Delta)) &= \ln(S(t)) + r + \tilde{\mu} h(t + \Delta) + \sqrt{h(t + \Delta)} z(t + \Delta), \quad (1.10a) \\
h(t + \Delta) &= \tilde{w} + \tilde{b}_1 h(t) + \tilde{b}_2 h(t - \Delta) + v_1(t) + v_2(t - \Delta), \quad (1.10b)
\end{align*}
\]

where

\[
\begin{align*}
\tilde{\mu} &= \mu + \eta^{-1}, \quad (1.11a) \\
\tilde{w} &= w + a_1 \eta^4 + a_2 \eta^4, \quad (1.11b) \\
\tilde{b}_i &= b_i + c_i / \eta^2 + a_i \eta^2, \quad (1.11c) \\
z(t) &= \eta y(t) - h(t)/\eta \sqrt{h(t)} \quad (1.11d) \\
v_i(t) &= c_i y(t) + a_i h(t)^2/y(t) - c_i h(t)/\eta^2 - a_i \eta^2 h(t) - a_i \eta^4. \quad (1.11e)
\end{align*}
\]

Note that by incorporating the lagged return innovation \(y(t - \Delta)\) and the reciprocals \(1/y(t)\) and \(1/y(t - \Delta)\), the change in variance is imperfectly correlated with the return.

\(^6\)Alternatively, one can define inverse of degree of freedom, similar to Bollerslev (1987), and the distribution converges to normal when the inverse of degree of freedom approaches to zero.
The conditional means of return and variance are given by

\[
E_t[\ln(S(t+\Delta)/S(t))] = r + \mu h(t+\Delta),
\]
\[
E_t[h(t+2\Delta)] = \bar{w} + \bar{b}_1 h(t+\Delta) + c_2 y(t) + a_2 h(t)^2/y(t).
\]

We now transform the IG-GARCH(2,2) into a component model that nests Christoffersen, Jacobs, Ornthanalai and Wang (2008). Define the long-run component \( q(t) \) of the variance process \((1.10b) \) as

\[
q(t) = -\rho_1 \tilde{w} \frac{1}{(1-\rho_1)(\rho_2-\rho_1)} + \frac{\rho_2}{\rho_2-\rho_1} h(t) + \frac{\tilde{b}_2}{\rho_2-\rho_1} h(t-\Delta) + \frac{1}{\rho_2-\rho_1} \nu_2(t-\Delta),
\]

where \( \nu_2(t) \) is given by \((1.11c) \), and where \( \rho_1 \) and \( \rho_2 \) are the respective smaller and larger roots of the quadratic equation

\[
\rho^2 - \tilde{b}_1 \rho - \tilde{b}_2 = 0,
\]

which are the eigenvalues of the transition equation \((1.9b) \). The short-run component is the deviation of variance from its long-run mean, \( h(t) - q(t) \). Substituting these into the IG-GARCH(2,2) dynamics \((1.9a)-(1.9b) \) yields the IG-GARCH component model which we will denote IG-GARCH(C) below

\[
\ln(S(t+\Delta)) = \ln(S(t)) + r + \mu h(t+\Delta) + \eta y(t+\Delta),
\]
\[
h(t+\Delta) = q(t+\Delta) + \rho_1 (h(t) - q(t)) + \nu_h(t),
\]
\[
q(t+\Delta) = w_q + \rho_2 q(t) + \nu_q(t),
\]

or equivalently,

\[
q(t+\Delta) = \sigma^2 + \rho_2 (q(t) - \sigma^2) + \nu_q(t),
\]

where \( \sigma^2 \) is the unconditional variance, and

\[
\sigma^2 = \frac{\tilde{w}}{1-\rho_1(1-\rho_2)}, \quad w_q = \frac{\tilde{w}}{1-\rho_1},
\]
\[
a_h = -\frac{\rho_1}{\rho_2-\rho_1} a_1 - \frac{1}{\rho_2-\rho_1} a_2, \quad a_q = \frac{\rho_2}{\rho_2-\rho_1} a_1 + \frac{1}{\rho_2-\rho_1} a_2
\]
\[
c_h = -\frac{\rho_1}{\rho_2-\rho_1} c_1 - \frac{1}{\rho_2-\rho_1} c_2, \quad c_q = \frac{\rho_2}{\rho_2-\rho_1} c_1 + \frac{1}{\rho_2-\rho_1} c_2
\]
\[
v_i(t) = c_i y(t) + a_i h(t)^2/y(t) - c_i h(t)/\eta^2 - a_i h(t) - a_i \eta t.
\]

The unit root condition, \( \rho_2 = 1 \), corresponds to the restriction \( \tilde{b}_2 = 1 - \tilde{b}_1 \). The expression for \( \sigma^2 \) shows that total
variance persistence in the component model is simply

\[ 1 - (1 - \rho_1)(1 - \rho_2) = \rho_2 + \rho_1(1 - \rho_2). \]

The component parameters can also be inverted to get the IG-GARCH(2,2) parameters

\[ a_1 = a_h + a_q \quad a_2 = -\rho_2 a_h - \rho_1 a_q \]
\[ b_1 = \rho_1 + \rho_2 \quad b_2 = -\rho_1 \rho_2 \]
\[ c_1 = c_h + c_q \quad c_2 = -\rho_2 c_h - \rho_1 c_q \]

This proves that the IG-GARCH(2,2) model is equivalent to the component model (1.14a)-(1.14c). In the IG-GARCH(1,1) special case studied in Christoffersen, Heston and Jacobs (2006), the long-run component in (1.14c) is effectively removed from the return dynamics.

1.2.5 The Gaussian Limit of the IG Model

We now show formally how the Gaussian models are nested by the Inverse Gaussian models. Consider the normalization of the innovation to the return process in (1.9a),

\[ z(t) = \frac{\eta y(t) - h(t)/\eta}{\sqrt{h(t)}}. \] (1.15)

This normalized Inverse Gaussian innovation converges to a Gaussian distribution as the degrees of freedom, \( h(t)/\eta^2 \), approach infinity. If we fix \( z(t) \) and \( h(t) \), and take the limit as \( \eta \) approaches zero, then the IG-GARCH(2,2) process (1.10a)-(1.10b) converges weakly to the Heston-Nandi (2000) GARCH(2,2) process in (1.2a):

\[
\begin{align*}
\text{ln}(S(t + \Delta)) & = \text{ln}(S(t)) + r + \mu h(t + \Delta) + \sqrt{h(t + \Delta)} z(t + \Delta), \\
h(t + \Delta) & = \omega + \beta_1 h(t) + \beta_2 h(t - \Delta) \\
& \quad + \alpha_1 (z(t) - \gamma_1 \sqrt{h(t)})^2 + \alpha_2 (z(t - \Delta) - \gamma_2 \sqrt{h(t - \Delta)})^2,
\end{align*}
\]
where the limit is taken as follows

\[ \tilde{w} = \omega - \alpha_1 - \alpha_2, \]
\[ a_i = \alpha_i / \eta^4, \]
\[ b_i = \beta_i + \alpha_i \gamma_i^2 + 2 \alpha_i \gamma_i / \eta - 2 \alpha_i / \eta^2, \]
\[ c_i = \alpha_i (1 - 2 \eta \gamma_i). \]

Written in component form, the limit is given by (1.3a)-(1.3b).

Our Inverse Gaussian Component model in (1.14a)-(1.14c) thus corresponds in the limit to the component model of Christoffersen, Jacobs, Ornthanalai, and Wang (2008). Christoffersen, Heston, and Jacobs (2006) show that the Inverse Gaussian GARCH(1,1) model in (1.4a)-(1.4b) nests the Heston-Nandi (2000) Gaussian GARCH(1,1) model in (1.1a)-(1.1b).

### 1.3 The Risk-Neutral Model and Option Valuation

To value options, we introduce the pricing kernel and the resulting risk-neutral dynamics. We then elaborate on the relationships between the risk-neutral and physical parameters. We first discuss the risk-neutralization for the most general IG-GARCH(2,2) process. Subsequently we discuss special cases nested by the most general specification.

#### 1.3.1 Risk-Neutralization

For the purpose of option valuation we need to derive the risk-neutral dynamics from the physical dynamics and pricing kernel. Risk-neutralization is more complicated for the Inverse Gaussian distribution than for the Gaussian distribution. We implement a volatility-dependent pricing kernel following Christoffersen, Heston, and Jacobs (2013), where

\[ M(t + \Delta) = M(t) \left( \frac{S(t + \Delta)}{S(t)} \right)^{\delta} \exp(\delta_0 + \delta_1 h(t + \Delta) + \xi h(t + 2\Delta)). \]  

(1.18)

Recent evidence by Cuesdeanu and Jackwerth (2015) suggests that the pricing kernel may be a non-monotonic function of returns. Accordingly, Christoffersen, Heston, and Jacobs (2013) show that in a GARCH framework, the variance-dependent log-kernel is a nonlinear and non-monotonic function of the path of spot returns. Henceforth we refer to it as the non-monotonic pricing kernel. If \( \xi > 0 \), the pricing kernel is U-shaped in returns. In Appendix A we show that the risk-free and the risky assets both satisfy the martingale restriction under the pricing kernel in
In Appendix B we show that the scaled return innovation $s_y y(t)$ is distributed Inverse Gaussian under the risk-neutral measure with variance $s_h h(t)$, where

$$s_y = 1 - 2c_1 \xi - 2\eta \phi,$$
$$s_h = \sqrt{1 - 2a_1 \xi \eta^4 s_y^{-3/2}}. \quad (1.19)$$

Inserting these definitions into the IG-GARCH(2,2) dynamics in (1.9) yields the risk-neutral process

$$\ln(S(t + \Delta)) = \ln(S(t)) + r + \mu^* h^* (t + \Delta) + \eta^* y^* (t + \Delta),$$
$$h^*(t + \Delta) = w^* + b_1 h^*(t) + b_2 h^*(t - \Delta) + c_1^* y(t) + c_2^* y(t - \Delta)$$
$$+ a_1^* h^*(t)^2 / y^*(t) + a_2^* h^*(t - \Delta)^2 / y^*(t - \Delta),$$

where

$$h^*(t) = s_h h(t), \quad y^*(t) = s_y y(t), \quad (1.21a)$$
$$\mu^* = \mu / s_h, \quad \eta^* = \eta / s_y, \quad w^* = s_h w, \quad (1.21b)$$
$$a_i^* = s_y a_i / s_h, \quad c_i^* = s_h c_i / s_y. \quad (1.21c)$$

The risk-neutral return process is IG-GARCH because the innovation $y^*(t + \Delta)$ has an Inverse Gaussian distribution under the risk-neutral measure. Notice that $b_1$ and $b_2$ are identical in the physical and risk-neutral processes. The risk-neutral process can also be written as a component model, the details are in Appendix C.

### 1.3.2 Preference Parameters and Risk-Neutral Parameters

Note that the risk-neutralization is specified for convenience in terms of the two reduced-form preference parameters $s_h$ and $s_y$. It is worth emphasizing that in fact only one extra parameter is required to convert physical to risk-neutral parameters. The martingale restriction for the risk-neutral dynamics is given by

$$\mu^* = \frac{\sqrt{1 - 2\eta_{\ast}^2} - 1}{\eta_{\ast}^2}, \quad (1.22)$$

We are grateful to our EFA discussant Fulvio Pegoraro for helping us clarify this derivation.
This imposes the following restriction between the physical parameters $\mu$ and the preference parameters $\phi$ and $\xi$

$$\mu = s_h \frac{\sqrt{1 - 2\eta/s_y} - 1}{\eta^2/s_y^2} = \sqrt{1 - 2a_1\xi\eta^4} \frac{\sqrt{1 - 2c_1\xi - 2\eta\phi - 2\eta - \sqrt{1 - 2c_1\xi - 2\eta\phi}}}{\eta^2}.$$  (1.23)

Given the physical parameters and the value of $\xi$ (or $s_y$), we can thus recover the value of the risk aversion parameter $\phi$ (or $s_h$). In other words, it takes only one additional parameter to convert between physical and risk-neutral parameters. To see this, alternatively re-write these restrictions as

$$s_y = \frac{(\frac{1}{2}\mu^2\eta^4 + (1 - 2a_1\xi\eta^4)\eta)^2}{(1 - 2a_1\xi\eta^4)\mu^2\eta^4}.$$  (1.24)

$$s_h = \frac{\mu\eta^2}{s_y^2(\sqrt{1 - 2\eta/s_y} - 1)}.$$  (1.25)

Because $s_h$ is now a function only of $s_y$ and physical parameters, this demonstrates that we can write (1.21a)–(1.21c) as a function of the physical parameters and one additional parameter, either the reduced form parameter $s_y$ or the preference parameter $\xi$.

### 1.3.3 Nested Option Models

The full risk-neutral valuation model has two components with inverse-Gaussian innovations. This model contains a number of simpler models as special cases. First consider the Gaussian limit of the risk-neutral dynamic. In the limit, as $\eta$ approaches zero, $\tilde{\mu}^* = \mu^* + \eta^{* - 1}$ approaches $-\frac{1}{2}$. Also in this limit, $s_h = s_y^{-1}$ as seen from equation (1.19). The risk-neutral process therefore converges to

$$\ln(S(t + \Delta)) = \ln(S(t)) + r - \frac{1}{2} h^*(t + \Delta) + \sqrt{h^*(t + \Delta)} z^*(t + \Delta),$$

$$h^*(t + \Delta) = \omega^* + \beta_1 h^*(t) + \beta_2 h^*(t - \Delta)$$

$$+ \alpha_1^* (z(t) - \gamma_1^* \sqrt{h^*(t)})^2 + \alpha_2^* (z(t - \Delta) - \gamma_2^* \sqrt{h^*(t - \Delta)})^2,$$

where

$$z^*(t + \Delta) = \frac{z(t + \Delta)}{\sqrt{s_h}} + \left( \frac{\tilde{\mu}}{\sqrt{s_h}} + \frac{\sqrt{s_y}}{2} \right) \sqrt{h(t + \Delta)},$$

$$\omega^* = s_h \omega, \quad \alpha_i^* = s_h^2 \alpha_i, \quad \gamma_i^* = \gamma_i + \tilde{\mu} s_h^{-1} + \frac{1}{2}. $$

This is the GARCH$(2,2)$ generalization of the risk-neutral version of the Gaussian GARCH$(1,1)$ model studied in Christoffersen, Heston, and Jacobs (2013). Following our previous analysis in equation (1.13), one may
alternatively express this as the risk-neutral Gaussian component model.

Further setting $\xi = 0$, or equivalently $s_h = 1$, we retrieve the GARCH(2,2) version of the Heston-Nandi (2000) model.

Finally, the risk-neutral versions of the GARCH(1,1) models are obtained straightforwardly by setting the appropriate parameters to zero, similar to the restrictions for the physical dynamics discussed in Section 1.2.

### 1.3.4 Option Valuation

Option valuation with this model is straightforward. Following Heston and Nandi (2000), the value of a European call option at time $t$ with strike price $X$ maturing at $T$ is equal to

$$
Call(S(t), h(t + \Delta), X, T) = S(t) \left( \frac{1}{2} + \frac{\exp^{-r(T-t)}}{\pi} \int_0^\infty \text{Re} \left[ \frac{X^{-i\varphi}g^*_t(i\varphi + 1, T)}{i\varphi S(t)} \right] d\varphi \right) - X \exp^{-r(T-t)} \left( \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \text{Re} \left[ \frac{X^{-i\varphi}g^*_t(i\varphi, T)}{i\varphi} \right] d\varphi \right),
$$

where $g^*_t(\varphi, T)$ is the conditional generating function for the risk-neutral process in (1.20). The conditional generating function $g_t(\varphi, T)$ under the physical measure is given by:

$$
g_t(\varphi, T) = E_t[S(T)^\varphi] = S(t)^\varphi \exp(A(t) + B(t)h(t + \Delta) + C(t)q(t + \Delta)),
$$

where

$$
A(T) = B(T) = C(T) = 0,
$$

$$
A(t) = A(t + \Delta) + \varphi r + (w_q - a_h\eta^4 - a_q\eta^4)B(t + \Delta) + (w_q - a_q\eta^4)C(t + \Delta) - \frac{1}{2} \ln(1 - 2(a_h + a_q)\eta^4 B(t + \Delta) - 2a_q\eta^4 C(t + \Delta)),
$$

$$
B(t) = \varphi\mu + (\rho_1 - (c_h + c_q)\eta^2 - (a_h + a_q)\eta^2)B(t + \Delta) - (c_q\eta^{-2} + a_q\eta^2)C(t + \Delta) + \eta^{-2} \frac{\sqrt{(1 - 2(a_q + a_h)\eta^4 B(t + \Delta) - 2a_q\eta^4 C(t + \Delta))(1 - 2\eta\varphi - 2(c_q + c_h)B(t + \Delta) - 2c_qC(t + \Delta))}}{\eta^2}
$$

$$
C(t) = (\rho_2 - \rho_1)B(t + \Delta) + \rho_2C(t + \Delta).
$$

This recursive definition requires computing equations (1.31-1.33) period-by-period with the terminal condition...
in (1.30) and then integrating \( g_t(\varphi, T) \) as in (1.28). Note that in equations (1.31)-(1.33) risk-neutral parameters should be used when valuing options. Note also that we have supplied the conditional generating function for the IG-GARCH(C) model. The corresponding functions for the nested models can be obtained as special cases of this function using the results above. Put options can be valued using put-call parity.

Armed with the formulas for computing option values, we are now ready to embark on an empirical investigation of our model.

1.4 Data and Estimation

1.4.1 Data

Our empirical analysis uses out-of-the-money S&P500 call and put options for the January 10, 1996 through December 26, 2012 period with a maturity between 14 and 365 days. We apply the filters proposed by Bakshi, Cao, and Chen (1997) as well as other consistency checks. Rather than using a short time series of daily option data, we use an extended time period, but we select option contracts for one day per week only. This choice is motivated by two constraints. On the one hand, it is important to use as long a time period as possible, in order to be able to identify key aspects of the model including volatility persistence. See for instance Broadie, Chernov, and Johannes (2007) for a discussion. On the other hand, despite the numerical efficiency of our model, the optimization problems we conduct are very time-intensive, because we use very large panels of option contracts. Selecting one day per week over a long time period is therefore a useful compromise. We use Wednesday data, because it is the day of the week least likely to be a holiday. It is also less likely than other days such as Monday and Friday to be affected by day-of-the-week effects. Moreover, following the work of Dumas, Fleming and Whaley (1998) and Heston and Nandi (2000), several studies have used a long time series of Wednesday contracts. The first Wednesday available in the OptionMetrics database is January 10, 1996, and so our sample is January 10, 1996 through December 26, 2012.

Panel A in Table 1.1 presents descriptive statistics for the return sample. The return sample is constructed from the S&P500 index returns. The return sample dates from January 2, 1990 through December 31, 2012. The standard deviation of returns, at 18.61\%, is substantially smaller than the average option-implied volatility, at 22.47\%. The higher moments of the return sample are consistent with return data in most historical time periods, with a small negative skewness and substantial excess kurtosis. Table 1.1 also presents descriptive statistics for the return sample from January 10, 1996 through December 26, 2012, which matches the option sample. In comparison
to the 1990-2012 sample, the standard deviation is somewhat higher, and average returns are somewhat lower. Average skewness and kurtosis in 1996-2012 are quite similar to the 1990-2012 sample.

Panels B and C of Table 1.1 present descriptive statistics for the option data by moneyness and maturity. Moneyness is defined as the implied futures price $F$ divided by strike price $X$. When $F/X$ is smaller than one, the contract is an out-of-the-money (OTM) call, and when $F/X$ is larger than one, the contract is an OTM put. The out-of-the-money put prices are converted into call prices using put-call parity. The sample includes a total of 29,022 option contracts with an average mid-price of 41.63 and average implied volatility of 22.47% as noted above. The implied volatility is largest for the OTM put options in Panel B, reflecting the well-known volatility smirk in index options. The implied volatility term structure in Panel C is roughly flat during the sample period.

1.4.2 Estimation

We now present a detailed empirical investigation of the model outlined in Section 1.2. We can separately evaluate the model’s ability to describe return dynamics and to fit option prices. But the model’s ability to capture the differences between the physical and risk-neutral distributions requires fitting both return and option data using the same, internally consistent, set of parameters.

We first use an estimation exercise that fits options and returns separately. We also employ sequential estimation following Broadie, Chernov, and Johannes (2007), who first estimate each model on returns only and subsequently assess the fit of each model to option prices in a second step where only risk-premium parameters are estimated. This procedure is also used by Christoffersen, Heston, and Jacobs (2013) in the context of a Gaussian GARCH(1,1) model with a quadratic pricing kernel. The consistency between returns and option contracts is the main reason behind the choice of sequential estimation. Bates (2000) notes that risk-neutral estimates are mostly inconsistent with the physical properties of return and variance time-series. Reconciliation between physical and risk-neutral measures is established with the sequential estimation since it is consistent with the returns while searching for optimal parameters that drive the risk-neutral process.

The advantage of sequential estimation is that it does not have any subjective choice of weights for return and option fit. Whenever we have two different optimization criteria (one for returns and one for options), the need for a weight arises and literature has many examples of weights. However, sequential estimation does not need a weighting scheme, and it avoids the subjective assignment of weights. Furthermore, given the estimates from time-series of returns, it reduces the number of parameters to be estimated. Thus, it simplifies the estimation procedure and decreases the chance of local optima. Although sequential estimation comes at a cost of higher
pricing errors in option contracts (since it is fully consistent with return time-series and has only one parameter to estimate), all the models are treated equally (one extra parameter to estimate).

Regarding the estimations, first consider returns. In the Inverse Gaussian case, the conditional density of the daily return is

\[
f(R(t)|h(t)) = \frac{\eta^{-3}}{\sqrt{2\pi}} \frac{h(t)}{(R(t) - r - \mu h(t))^3} \times \exp\left(-\frac{1}{2} \left( \frac{R(t) - r - \mu h(t)}{\eta} - \frac{h(t)}{\eta^2} \sqrt{\frac{\eta}{R(t) - r - \mu h(t)}} \right)^2 \right).
\]

The return log-likelihood is summed over all return dates.

\[
\ln L^R \propto \sum_{t=1}^{T} \{ \ln(f(R(t)|h(t))) \}.
\]  

(1.34)

We can therefore obtain the physical parameters \( \Theta \) by estimating

\[
\Theta_{Return} = \arg \max_{\Theta} \ln L^R.
\]  

(1.35)

Now consider the options data. Define the Black-Scholes Vega (BSV) weighted option valuation errors as

\[
\varepsilon_i = \left( Call^Mkt_i - Call^Mod_i \right) / BSV^Mkt_i,
\]

where \( Call^Mkt_i \) represents the market price of the \( i \)th option, \( Call^Mod_i \) represents the model price, and \( BSV^Mkt_i \) represents the Black-Scholes vega of the option (the derivative with respect to volatility) at the market implied level of volatility. Assume these disturbances are i.i.d. normal so that the option log-likelihood is

\[
\ln L^O \propto -\frac{1}{2} \sum_{i=1}^{N} \{ \ln(s^2_i) + \varepsilon_i^2/s^2_i \}.
\]  

(1.36)

where we can concentrate out \( s^2_i \) using the sample analogue \( \hat{s}^2 = \frac{1}{N} \sum_{i=1}^{N} \varepsilon_i^2 \). We use the term structure of interest rates from OptionMetrics when pricing options.

The vega-weighted option errors are very useful because it can be shown that they are an approximation to implied volatility based errors, which have desirable statistical properties. Unlike implied volatility errors, they do not require Black-Scholes inversion of model prices at every step in the optimization, which is very costly in large
scale empirical estimation exercises such as ours. We obtain the risk-neutral parameters $\Theta^*$ based on options data by estimating

$$
\Theta^*_\text{Option} = \arg \max_{\Theta^*} \ln L^O.
$$

(1.37)

Note that both estimation exercises mentioned above ignore the specification of the pricing kernel, and are therefore uninformative about the choice between the log-linear and non-monotonic pricing kernels. We thus conduct a third estimation exercise where we sequentially estimate the non-monotonic pricing kernel parameter, $\xi$, on options only, keeping all the physical parameters from (1.35) fixed. We thus estimate

$$
\xi_{\text{Seq}} = \arg \max_{\xi} \ln L^O.
$$

(1.38)

Sequential estimation is of course only conducted for the models with non-monotonic pricing kernels. Our sequential estimation approach follows that in Broadie, Chernov, and Johannes (2007) and Christoffersen, Heston, and Jacobs (2013).

### 1.5 Empirical Results

Because our specification nests several models, it allows for a comparison of the relative importance of model features. Specifically, we can compare the contribution of a second stochastic volatility factor, fat-tailed innovations, and a non-monotonic (or variance-dependent) pricing kernel. We can quantify the contribution of these features in separately explaining the time series of returns and the cross-section of option prices, as well as in explaining returns and options together, which we do in a sequential estimation exercise.

While a horserace based on model fit is of interest, it is also relevant to verify whether the different model features are complements rather than substitutes. In theory this should be the case: the second volatility factor should improve the modeling of the term structure of volatility, and therefore the valuation of options of different maturities, especially long-maturity options. In contrast, the fat-tailed IG innovation should prove most useful to capture the moneyness dimension for short-maturity out-of-the-money options, which is usually referred to as the smirk. The non-monotonic pricing kernel has an entirely different purpose, because its relevance lies in the joint modeling of index returns and options, rather than the modeling of options alone.

Tables 1.2-1.6 present the empirical results. Table 1.2 presents estimation results using returns data. The results include parameter estimates and log-likelihoods, as well as several implications of the parameter estimates such

⁸See for instance Carr and Wu (2007) and Trolle and Schwartz (2009) for applications of $BSV^{Mkt}$ weighted option errors.
as moments and persistence. Table 1.3 presents similar results for the estimation based on option data, and Table 1.4 does the same for the sequential estimation based first on returns and subsequently on options. Table 1.4 also reports the improvement in fit for the non-monotonic pricing kernel over the linear pricing kernel in terms of log-likelihood values. Tables 1.5 and 1.6 provide more details on the models’ fit across moneyness and maturity categories for the three estimation exercises in Tables 1.2 and 1.4.

1.5.1 Fitting Returns and Fitting Options

We will organize our initial discussion around the measures of fit (i.e. log-likelihood values) for the different models contained in Table 1.2 (return fitting) and Table 1.3 (option fitting). We have results for the fit of six models in these tables. Of these six models, three have Gaussian innovations and three are characterized by fat-tailed Inverse Gaussian innovations. Two models have two variance factors, two have one factor. For comparison we also estimate two models that have no variance dynamics, which we refer to as homoskedastic models.

The most highly parameterized two-factor model with fat tails fits the returns and options data best, as can be seen in Tables 1.2 and 1.3 while the most restrictive single factor Gaussian model fits worst, which is not surprising in an in-sample exercise.

All the two-factor models have substantially higher likelihood values than all the one-factor models. The two-factor models have three more parameters than the corresponding one-factor models, and two times the difference in the log-likelihoods is asymptotically distributed chi-square with three degrees of freedom. The 99.9% p-level for this test is 16.3. In the case of the option-based estimation in Table 1.3 the improvement provided by the second factor is very large, with a likelihood improvement of approximately 5,000. This suggests that the most important feature in accurately modelling option prices is the correct specification of the volatility dynamics. This finding confirms the results in Andersen, Fusari, and Todorov (2015), where the biggest improvement in fit also stems from a second factor.

The inclusion of a second factor also significantly improves the return fit in Table 1.2. For example, for the Gaussian case, twice the difference in the log-likelihood between the two-factor and one-factor models is 100, and for the fat-tailed case the corresponding number is 71. These test statistics are highly significant. We conclude that a second factor is important in describing the underlying returns as well as option prices.

When comparing IG versus Gaussian models, Tables 1.2 and 1.3 show that adding the single parameter \( \eta \) in the IG models increases the return and option likelihoods substantially. In Table 1.3 the likelihood improvements

---

9This could also be tested using the Model Confidence Set approach of Hansen, Lunde, and Nason (2011)
are again in the thousands. The improvements in the return likelihoods in Table 1.2 are less dramatic but still statistically significant at conventional confidence levels.

Notice also that the magnitude of improvement contributed by the fat-tailed IG feature depends on the other model features. For the option-based estimation in Table 1.3 the improvement in option log-likelihood is 4,314 for the homoskedastic case, 1,853 for the one-factor GARCH, and 1,575 for the two-factor GARCH. For the return-based estimation in Table 1.2 the improvements in log-likelihood are 3.5 for the homoskedastic case, 48.1 for the one-factor GARCH, and 33.6 for the two-factor GARCH. The IG feature improves the fit more in the case of the simpler one-factor model than in the case of the two-factor model. Non-normal innovations and a second volatility component are therefore to some extent substitutes in model specification.

1.5.2 Sequential Estimation of the Non-Monotonic Pricing Kernel Parameter

Table 1.2 contains return-based estimates of the physical distributions. Table 1.3 contains option-based estimates of the risk-neutral distribution. Neither table is informative about the pricing kernel. In Table 1.4 we therefore use the physical parameter estimates from Table 1.2 and estimate only the variance-dependent pricing kernel parameter $\xi$ by fitting options. Table 1.4 reports risk-neutral values of all parameters, but only $\xi$ is estimated from options.

The penultimate column in Panel B of Table 1.4 reports the option likelihoods for the four dynamic models with variance-dependent pricing kernel. The last column in Panel B shows the difference between the option likelihood for optimal $\xi$ and that for $\xi = 0$, where the options are valued using the risk-neutralized parameters derived from Table 1.2 with the linear pricing kernel (since $\xi = 0$, there is no variance-dependency and the pricing kernel is linear in returns).

The increase in option log-likelihood when allowing for a variance-dependent (i.e. non-monotonic in returns) pricing kernel and adding just a single parameter is again in the thousands.

Table 1.4 shows that the log-likelihood increase due to the more general pricing kernel is 6,644 in the single factor Gaussian model, and 9,548 in the corresponding Inverse Gaussian model. In case of the two-factor models, the improvements are even higher: The non-monotonic kernel improves the two-factor likelihoods by 9,336 in the Gaussian model and 11,180 in the Inverse Gaussian model.

We conclude that the importance of modeling a more general pricing kernel depends on the models’ ability to capture the tails of the distribution. The richer dynamics of two-factor models allow them to better fit the fat tails, and a non-monotonic pricing kernel captures this property by allowing the model’s physical parameters to fit the returns and risk-neutral parameters to fit options in the same model. Complex modeling of risk premia
complements adequate modeling of return dynamics.

Table [1.4] is also interesting in that it shows that the two key conclusions from Tables [1.2] and [1.3] still obtain: Allowing for Inverse Gaussian innovations improves the fit, as does allowing for a second variance component. Note that in Table [1.4] these conclusions are based on option fit but use return-based estimates, which shows that these findings are not merely in-sample phenomena.

Figure [1.1] complements Table [1.4] by plotting the implied volatility RMSE percentages (top panel) and log-likelihood values (bottom panel) for different values of the $\xi$ parameter in the models we consider. Figure [1.1] shows that the IG-GARCH component model we propose has lower RMSE and higher log-likelihood values for the optimal $\xi$ parameter and indeed for a wide range of values around the optimum. The linear pricing kernel corresponds to the left-most point on the curves where $\xi = 0$.

### 1.5.3 Capturing Dynamics in Higher Moments

Examination of the parameter estimates in Tables [1.2][1.4] reveals the main reason for the superior performance of the two-factor models. For the returns-based estimation in Table [1.2] the persistence of the single factor estimates is 0.98 at a daily frequency for the Gaussian and the Inverse Gaussian model. For the two-factor models, the long-run factor is always very persistent ($\rho_2$ is around 0.99), but the persistence of the short-run factor, $\rho_1$, is 0.71 in the Gaussian model and 0.74 in the Inverse Gaussian model. The single-factor models are forced to compromise between slow and fast mean reversion, leading to a deterioration in fit in some parts of the sample.

Figures [1.2] and [1.3] provide additional perspective on the differences between the GARCH(1,1) and component models. Figure [1.2] plots the spot variance for all models using the return-based estimates. Figure [1.3] also uses the return-based estimates to plot conditional (“leverage”) correlation between returns and variance, $\text{Corr}_t[R(t + \Delta), h(t + 2\Delta)]$, which is informative about the third moment dynamics, and conditional standard deviation of variance, $\sqrt{\text{Var}_t[h(t + 2\Delta)]}$, which is informative about the fourth moment dynamics. The formulas used for these conditional moments are contained in Appendix D.

In Figure [1.2] we can see that component model total variance (i.e. $h(t)$) is more variable and has the ability to increase faster than the GARCH(1,1), thanks to its short-run component (i.e. $h(t) - q(t)$). During the recent financial crisis the variances in the component models jump to a higher level than do the GARCH(1,1) variances. Consistent with this finding, the conditional standard deviation of variance (conditional correlation between returns and variance) of the component models in Figure [1.3] is higher in level (more negative) and more noisy than those of GARCH(1,1) models.
Figure 1.4 graphs the term structure of variance, skewness and kurtosis using the derivatives of the moment generating function. Variance, skewness and kurtosis are defined by

\[
Var_t(T) = \frac{\partial^2 \ln g_t(\varphi, T)}{\partial \varphi^2}|_{\varphi=0},
\]

(1.39)

\[
Skew_t(T) = \frac{\partial^3 \ln g_t(\varphi, T)}{\partial \varphi^3}|_{\varphi=0} \left( \frac{\partial^2 \ln g_t(\varphi, T)}{\partial \varphi^2}|_{\varphi=0} \right)^{3/2},
\]

(1.40)

\[
Kurt_t(T) = \frac{\partial^4 \ln g_t(\varphi, T)}{\partial \varphi^4}|_{\varphi=0} \left( \frac{\partial^2 \ln g_t(\varphi, T)}{\partial \varphi^2}|_{\varphi=0} \right)^2 - 3.
\]

(1.41)

The plots in the first column of Figure 1.4 show variance normalized by unconditional variance of each model, the second column shows skewness and the third column shows kurtosis. Each row corresponds to a different model. The initial variance is set to twice the unconditional model variance in the solid lines and the initial variance is set to one-half the unconditional variance in the dashed lines. For the component models we set the long-run variance component, \(q_t\) equal to three-quarters of total variance, \(h(t)\). We use the return-based parameters in Table 1.2 to plot Figure 1.4.

The left-side panels in Figure 1.4 highlight the differences between the GARCH(1,1) and component models. The impact of the current conditions on the future variance is much larger for the component models, and this is of course due to the persistence of the long-run component. For the GARCH(1,1) model, the conditional variance converges much quicker to the long-run variance.

Figure 1.4 also shows that the term structures of skewness and kurtosis in the models differ between one-factor and component models. The one-factor models generate strongly hump-shaped term structures whereas the component models do so to a much lesser degree.

Figure 1.4 confirms that the Gaussian and Inverse Gaussian models do not differ much in the term structure dimension, and also indicates that the effects of shocks last much longer in the component models.

Figures 1.5 and 1.6 repeat Figures 1.2 and 1.3 but uses the option-based parameters in Table 1.3 rather than the physical parameters in Table 1.2. In Figure 1.5, the variance paths for the GARCH(1,1) and component models are very different compared to the return-implied paths in Figure 1.2. Note in particular that the short-run component in the component models strongly differs between Figures 1.2 and 1.5. In Figure 1.6 the time path of the conditional standard deviation of variance in the right-side panels is rather similar to the one from Figure 1.3 but this is not the case for the conditional correlation in the left-side panels.

Most model implications can be easily understood by inspecting the parameter estimates in Tables 1.2 and 1.4. In the case of the risk-neutral estimates from options in Table 1.4 a first important conclusion is that the component
models are more persistent than the GARCH(1,1) model, but the differences are smaller than in the case of the return-based estimates in Table 1.2. As a result, the impact of the current conditions on the future variance is larger for the component models, but the differences with the GARCH(1,1) model are larger for the return-based estimates. Second, results are always very similar for the Gaussian and Inverse Gaussian models, which is not surprising. Third, and most importantly, the risk-neutral dynamics are more persistent than physical dynamics. As a result, the impact of the current conditions on the future variance is much larger for the option-implied risk-neutral estimates, regardless of the model.

When estimating the models using returns and options sequentially in Table 1.4, the persistence of the models, and consequently the impact of the current conditions on the future variance, is close to the physical persistence based on returns in Table 1.2 since we fix the physical parameters in this estimation to the optimized returns-based parameter estimates.

1.5.4 The Relative Importance of Model Features for Option RMSE

We now perform an assessment of the relative importance of the three model features for option fitting. To this end consider the “All” RMSE in the last column of Table 1.5 which contains the implied volatility root mean squared error across all options. Panel A uses the return-based estimates from Table 1.2, Panel B uses the option-based estimates from Table 1.3, and Panel C uses the sequential estimates from Table 1.4.

The last column in Table 1.5 enables us to make six pairwise comparisons of GARCH(1,1) and component GARCH(C) models. The improvement from adding a second volatility factor ranges from 4.87% (1 − 5.0694/5.3289) and 3.8% in Panel A, to 16.24% and 15.45% in Panel B, and finally 9.71% and 6.6% in Panel C. On average the improvement from adding a second volatility factor is 9.45%. The improvement from adding a second volatility factor is largest in Panels B and C where the non-monotonic pricing kernel affects the results. The second volatility component and the variance-dependent pricing kernel thus appear to be complements.

The last column in Table 1.5 also enables us to compute six pairwise comparisons of GARCH versus IG-GARCH models. The IV-RMSE improvement from adding fat tails ranges from 2.12% and 1.02% in Panel A, to 6.18% and 5.29% in Panel B, and 7.38% and 4.20% in Panel C. The overall improvement from adding fat tails is 4.4% and thus considerably lower than from adding a second volatility factor. The improvement from adding fat tails is again largest in Panels B and C where the non-monotonic pricing kernel affects the results. Fat tails and a variance-dependent pricing kernel thus also appear to be complements rather than substitutes.

Finally, comparing Panels C and A in Table 1.5 allows us assess the importance of a variance-dependent versus
a linear pricing kernel. The improvement from allowing for a variance-dependent pricing kernel is $13.39\%$ $(1 - 4.6155/5.3289)$ for the GARCH(1,1) model, $18.05\%$ for the IG-GARCH(1,1) model, $17.80\%$ for the GARCH(C) model, and $20.44\%$ for the IG-GARCH(C) model. On average the improvement is $17.42\%$. The improvement from allowing for a variance-dependent pricing kernel is larger for IG than for Gaussian GARCH models, and it is larger for two-factor than for single-factor models which again suggests that the three features we investigate are complements rather than substitutes.

1.5.5 Capturing Smiles and Smirks

In Tables 1.5 and 1.5 we further investigate the model option fit across the moneyness and maturity categories defined in Table 1.1. Tables 1.5 and 1.6 report implied volatility RMSE and bias (in percent) by moneyness, and maturity, respectively.

Table 1.5 shows that the IG-GARCH(C) model we propose fits the data best in almost all moneyness categories. Not surprisingly, all models have most difficulty fitting the deep in-the-money calls (corresponding to deep out-of-the-money puts) which are very expensive. It is also not surprising that the fit in Panel B is almost always better than in Panel C which in turn is better than in Panel A. In Panel B, the option fit drives all the parameter estimates, in Panel C only $\xi$ is estimated on options, whereas in Panel A no parameters are fitted to option prices. Again, the most important conclusion from Table 1.5 is that the IG-GARCH(C) model performs well regardless of implementation and moneyness category.

Panel A of Table 1.5 also shows that the large RMSEs are largely driven by bias. The bias is defined as market IV less model IV. Positive numbers thus indicate that the model underprices options on average. Panel A shows that the models with linear pricing kernel estimated on returns only have large positive biases in every moneyness category. In Panel B, where all parameters are estimated on options, the bias is much closer to zero. In Panel C the bias is much smaller than in Panel A but it is still fairly large for deep in-the-money calls.

Table 1.6 reports the implied volatility RMSE and bias by maturity. The IG-GARCH component model now performs the best in all categories. Table 1.6 also shows that all models tend to underprice options (i.e. positive bias) at most maturities except for the very long-dated options.

Tables 1.5 and 1.6 indicate that the fat-tailed Inverse Gaussian distribution is also helpful in fitting the data. Fat-tailed innovations increase the values of short-term out-of-the-money options, whereas two-factor dynamics increase the tails and values of long-term out-of-the-money options. Tables 1.5 and 1.6 demonstrate that these model features are to some extent complementary.
The increases in likelihood due to fat-tailed innovations are much smaller than those due to the second volatility factor. This observation is consistent across estimation exercises and is confirmed by inspecting stylized facts. Figure 1.5 indicates that the variance paths are very similar for the models with Gaussian and Inverse Gaussian innovations for the option-based estimation results. However, this is unsurprising and not necessarily very relevant for the purpose of option valuation. Models with very similar variance paths can greatly differ with respect to their (conditional) third and fourth moments, and these model properties are of critical importance for option valuation, and for capturing smiles and smirks in particular. Therefore, we again look at conditional correlation and standard deviation of variance paths for the options-based estimations in Figure 1.6 which indicates substantial differences between the conditional correlation and standard deviation of variance paths for the Gaussian and Inverse Gaussian models. However, perhaps somewhat surprisingly, Figures 1.5 and 6 clearly indicate that the differences between the GARCH(1,1) and component models are actually larger than the differences between the Gaussian and Inverse Gaussian models in this dimension. This is surprising because a priori we expect the second factor to be more important for term structure modeling, as confirmed by Figure 1.4. The conditional moments in Figures 1.5 and 1.6 are more important for the modeling of smiles and smirks, and a priori we expect the modeling of the conditional innovation to be more important in this dimension. However, it seems that the second volatility factor is also of first-order importance in this dimension.

Figure 1.7 further illustrates the component model’s flexibility. We plot model-based implied volatility smiles using our proposed IG component model and the parameter values from Table 1.4. The total spot volatility, $\sqrt{h(t)}$, is fixed at 25% per year in all panels. In the top panel, the long run volatility factor, $\sqrt{q(t)}$, is set to 20%, in the middle panel it is set to 25%, and in the bottom the top panel it is set to 30%. We also show the IG-GARCH(1,1) model for reference. It is of course the same across the three panels. Figure 1.7 shows that the second volatility factor gives the model a great deal of flexibility in modeling the implied volatility smile.

### 1.5.6 Model-Implied Relative Risk Aversion

When using the standard log-linear pricing kernel, the coefficient of relative risk aversion is simply (the negative of) $\phi$. In the non-monotonic pricing kernel the computation of risk-aversion is slightly more involved and we therefore provide some discussion here.

Assume a representative agent with utility function $U(S(t))$ then the one-period coefficient of relative risk aversion can be written

$$RRA(t) \equiv -S(t) \frac{U''(S(t))}{U'(S(t))} = -S(t) \frac{M'(t)}{M(t)} = -S(t) \frac{\partial \ln (M(t))}{\partial S(t)},$$

(1.42)
where we have used the insight of Jackwerth (2000) to link risk aversion to the pricing kernel. From (1.18) we have that
\[
\frac{\partial \ln (M(t))}{\partial S(t)} = \frac{\phi}{S(t)} + \xi \frac{\partial h(t+\Delta)}{\partial S(t)}.
\] (1.43)

In the Gaussian model we have
\[
\frac{\partial h(t+\Delta)}{\partial S(t)} = \frac{\partial h(t+\Delta)}{\partial z(t)} \frac{\partial z(t)}{\partial S(t)} = \frac{2\alpha_1 (z(t) - \gamma_1 \sqrt{h(t)})}{\sqrt{h(t)}S(t)}.
\] (1.44)

Combining (1.43) and (1.44) we get a relative risk aversion of
\[
RRA(t) = -\phi - 2\alpha_1 \xi \frac{(z(t) - \gamma_1 \sqrt{h(t)})}{\sqrt{h(t)}}.
\]

Note as indicated above that the parameter $\phi$ does not in itself capture relative risk aversion unless $\xi = 0$ which corresponds to the linear pricing kernel.

Using the law of iterated expectations we can now compute the expected $RRA$ as
\[
E[RRA(t)] = -\phi + 2\alpha_1 \xi \gamma_1.
\]

Using the GARCH(1,1) parameter estimates in Tables [1.2] and [1.4] and the results in Appendix B of Christoffersen, Heston and Jacobs (2013), we get
\[
\phi = - (\bar{\mu} + \gamma_1) (1 - 2\alpha_1 \xi) + \gamma_1 - \frac{1}{2} \approx 20.26,
\]
so that we get
\[
E[RRA(t)] \approx -20.26 + 2\alpha_1 \xi \gamma_1 \approx 1.44.
\]

This result shows that the non-monotonic pricing kernel delivers reasonable coefficients of relative risk aversion, and furthermore that it is important not to rely on (the negative of) $\phi$ as a measure of $RRA$ when using the non-monotonic pricing kernel. Determining which equilibrium models are consistent with our pricing kernel is an interesting question that we leave for future work.
1.6 Conclusion

We find that multiple volatility factors, fat-tailed return innovations, and a variance-dependent pricing kernel all provide economically and statistically significant improvements in describing S&P500 returns and option prices. A variance-dependent pricing kernel is economically most important and improves option fit by 17% on average and more so for two-factor models. A second volatility factor improves the option fit by 9% on average. Fat tails improve option fit by just over 4% on average, and more so when a variance-dependent pricing kernel is applied. Our results show that overall these three features are complements rather than substitutes. This indicates that while proper specification of volatility dynamics is quantitatively most important in option models, the interdependent explanatory power of different features make it essential to evaluate them in a properly specified model that nests all of these features.
1.7 Figures

Figure 1.1: RMSE and Option Likelihood Values versus $\xi$.

Notes: We plot the RMSE (top panel) and the option likelihood function (bottom panel) as a function of the non-monotonic pricing kernel parameter, $\xi$. All other parameter values are fixed at their optimal values from Table 1.2.
Figure 1.2: Spot Variance Paths Using Return-Based Estimates

Notes: For each model we plot the spot variance components over time. The parameter values are obtained from MLE on returns in Table 1.2.
Figure 1.3: Leverage Correlation and Volatility of Variance Using Return-Based Estimates

Notes: For each model we plot the conditional correlation and the conditional standard deviation of variance. In the left panels, we plot the conditional correlation between return and variance as implied by the models. In the right panels, we plot the conditional standard deviation of conditional variance. The scales are identical across the rows of panels to facilitate comparison across models. The parameter values are obtained from MLE on returns in Table 1.2.
Figure 1.4: Term Structure of Variance, Skewness and Kurtosis

Notes: We plot the term structure of variance, skewness and excess kurtosis with high (solid) and low (dashed) initial variance for 1 through 250 trading days. Conditional variance is normalized by the unconditional variance, $\sigma^2$. For the low initial variance, the initial value of $q(t + \Delta)$ is set to $0.75\sigma^2$, and the initial value of $h(t + \Delta)$ is set to $0.5\sigma^2$. For the high initial variance, the initial value of $q(t + \Delta)$ is set to $1.75\sigma^2$, and the initial value of $h(t + \Delta)$ is set to $2\sigma^2$. The return-based parameter values from Table 1.2 are used.
Figure 1.5: Spot Variance Paths Using Option-Based Estimates

Notes: We plot the spot variance components over time. The parameter values are obtained from MLE on options in Table 1.3
Figure 1.6: Leverage Correlation and Volatility of Variance Using Option-Based Estimates

Notes: For each model we plot the conditional correlation and the conditional standard deviation of variance. In the left panels, we plot the conditional correlation between return and variance as implied by the models. In the right panels, we plot the conditional standard deviation of conditional variance. The scales are identical across the rows of panels to facilitate comparison across models. The parameter values are obtained from MLE on options in Table 1.3.
Figure 1.7: Model-Based Implied Volatility Smiles in IG-GARCH Component Model

Notes: We plot model-based implied volatility smiles for 30 days to maturity from the IG-GARCH(1,1) and IG-GARCH(C) models. Long-run volatility, $\sqrt{q(t)}$, is set to 20% (top panel), 25% (middle panel), and 30% (bottom panel). Total volatility, $\sqrt{h(t)}$ is set to 25% in all panels. The parameter estimates from Table 1.4 are used to generate the model prices. Model implied volatilities are calculated by inverting the Black-Scholes formula on the model prices.
1.8 Tables
Table 1.1: Returns and Options Data

Panel A. Return Characteristics (Annualized)

<table>
<thead>
<tr>
<th></th>
<th>1990-2012</th>
<th>1996-2012</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>6.06%</td>
<td>4.99%</td>
</tr>
<tr>
<td>Std. deviation</td>
<td>18.61%</td>
<td>20.57%</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.228</td>
<td>-0.217</td>
</tr>
<tr>
<td>Excess kurtosis</td>
<td>8.461</td>
<td>7.235</td>
</tr>
</tbody>
</table>

Panel B. Option Data by Moneyness

<table>
<thead>
<tr>
<th></th>
<th>F/X ≤ .80</th>
<th>.80 &lt; F/X ≤ .90</th>
<th>.90 &lt; F/X ≤ 1.00</th>
<th>1.00 &lt; F/X ≤ 1.10</th>
<th>1.10 &lt; F/X ≤ 1.20</th>
<th>F/X &gt; 1.20</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Contracts</td>
<td>720</td>
<td>3,819</td>
<td>8,413</td>
<td>8,033</td>
<td>5,778</td>
<td>2,259</td>
<td>29,022</td>
</tr>
<tr>
<td>Average IV</td>
<td>23.11%</td>
<td>19.65%</td>
<td>18.79%</td>
<td>22.09%</td>
<td>27.03%</td>
<td>30.52%</td>
<td>22.47%</td>
</tr>
<tr>
<td>Average Price</td>
<td>62.94</td>
<td>40.71</td>
<td>43.62</td>
<td>47.93</td>
<td>33.18</td>
<td>28.14</td>
<td>41.63</td>
</tr>
<tr>
<td>Average Spread</td>
<td>1.30</td>
<td>1.42</td>
<td>1.89</td>
<td>2.06</td>
<td>1.58</td>
<td>1.41</td>
<td>1.76</td>
</tr>
</tbody>
</table>

Panel C. Option Data by Maturity

<table>
<thead>
<tr>
<th></th>
<th>DTM ≤ 30</th>
<th>30 &lt; DTM ≤ 60</th>
<th>60 &lt; DTM ≤ 90</th>
<th>90 &lt; DTM ≤ 120</th>
<th>120 &lt; DTM ≤ 180</th>
<th>DTM &gt; 180</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Contracts</td>
<td>2,771</td>
<td>6,127</td>
<td>4,565</td>
<td>2,720</td>
<td>4,019</td>
<td>8,820</td>
<td>29,022</td>
</tr>
<tr>
<td>Average IV</td>
<td>24.93%</td>
<td>23.33%</td>
<td>22.45%</td>
<td>22.93%</td>
<td>21.74%</td>
<td>21.32%</td>
<td>22.47%</td>
</tr>
<tr>
<td>Average Price</td>
<td>18.06</td>
<td>26.83</td>
<td>31.88</td>
<td>39.29</td>
<td>48.54</td>
<td>61.92</td>
<td>41.63</td>
</tr>
<tr>
<td>Average Spread</td>
<td>0.94</td>
<td>1.31</td>
<td>1.59</td>
<td>1.87</td>
<td>1.97</td>
<td>2.29</td>
<td>1.76</td>
</tr>
</tbody>
</table>

We present descriptive statistics for daily return data from January 2, 1990 through December 31, 2012, as well as for daily return data from January 10, 1996 through December 26, 2012. We use Wednesday closing options contracts from January 10, 1996 through December 26, 2012.
Table 1.2: Maximum Likelihood Estimation Results for Return Distribution

Panel A: Parameter Estimates

<table>
<thead>
<tr>
<th>Gaussian Models</th>
<th>$\mu$</th>
<th>$\omega$</th>
<th>$\beta_1$</th>
<th>$\alpha_1$</th>
<th>$\gamma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homoskedastic</td>
<td>0.78</td>
<td>1.373E-04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1.12E+0)</td>
<td>(1.12E-6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| GARCH(1,1) | $\mu$  | $\omega$  | $\beta_1$  | $\alpha_1$  | $\gamma$  |
|            | 1.10   | -1.396E-06| 0.900      | 3.761E-06   | 145.7    |
| (1.13E+0)  | (1.35E-7) | (7.84E-3) | (2.30E-7) | (1.02E+1) |

| GARCH(C) | $\mu$  | $\omega$  | $\beta_1$  | $\alpha_1$  | $\gamma$  |
|          | 1.26   | 1.473E-06 | 0.705      | 9.979E-07   | 840.6    |
| (1.13E+0) | (1.53E-7) | (3.13E-2) | (4.32E-7) | (2.70E+6) |

<table>
<thead>
<tr>
<th>IG Models</th>
<th>$\mu$</th>
<th>$w$</th>
<th>$w$</th>
<th>$\beta_1$</th>
<th>$\alpha_1$</th>
<th>$\gamma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homoskedastic</td>
<td>0.78</td>
<td>1.372E-04</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1.23E+0)</td>
<td>(1.13E-6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| IG-GARCH(1,1) | $\mu$  | $w$  | $\beta_1$  | $\alpha_1$  | $\gamma$  |
|              | 1.16   | -1.469E-06| 21.82      | 3.190E+07   | 4.047E-06 |
| (1.13E+0)  | (1.28E-7) | (3.88E+0) | (1.03E+7) | (2.15E-7) |

| IG-GARCH(C) | $\mu$  | $w$  | $\beta_1$  | $\alpha_1$  | $\gamma$  |
|             | 1.23   | 1.393E-06 | 0.743      | 2.247E+06   | 6.987E-07 |
| (1.12E+0)  | (1.33E-7) | (3.00E-2) | (9.17E+6) | (4.81E-7) |

Panel B: Model Properties and Likelihoods

<table>
<thead>
<tr>
<th>Gaussian Models</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Homoskedastic</td>
<td>Return Mean</td>
<td>Annualized Volatility</td>
<td>Volatility Persistence</td>
<td>Uncond. persistence</td>
<td>Uncond. skewness</td>
</tr>
<tr>
<td></td>
<td>5.99%</td>
<td>18.60%</td>
<td>0.000</td>
<td>3.000</td>
<td></td>
</tr>
</tbody>
</table>

| GARCH(1,1) | Return Mean | Annualized Volatility | Volatility Persistence | Uncond. persistence | Uncond. skewness | Uncond. kurtosis | Log Likelihood |
|            | 6.48%       | 17.00%           | 0.979387             | 0.015              | 4.750            |                  | 18,781.4      |

| GARCH(C) | Return Mean | Annualized Volatility | Volatility Persistence | Uncond. persistence | Uncond. skewness | Uncond. kurtosis | Log Likelihood |
|          | 6.91%       | 16.91%           | 0.996170             | 0.024              | 5.199            |                  | 18,831.4      |

| IG Models | | | | | |
|-----------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| Homoskedastic | Return Mean | Annualized Volatility | Volatility Persistence | Uncond. persistence | Uncond. skewness | Uncond. kurtosis | Log Likelihood |
|             | 5.99%       | 18.59%           | -0.033             | 3.000             |                  |                  | 17,551.5      |

| IG-GARCH(1,1) | Return Mean | Annualized Volatility | Volatility Persistence | Uncond. persistence | Uncond. skewness | Uncond. kurtosis | Log Likelihood |
|               | 6.61%       | 16.92%           | 0.982695             | -0.152             | 4.775            |                  | 18,829.5      |

| IG-GARCH(C) | Return Mean | Annualized Volatility | Volatility Persistence | Uncond. persistence | Uncond. skewness | Uncond. kurtosis | Log Likelihood |
|             | 6.78%       | 16.81%           | 0.996802             | -0.099             | 5.247            |                  | 18,865.0      |

Parameter estimates are obtained by an MLE estimation on returns from January 2, 1990 through December 31, 2012. Data description can be found in Table 1. For each model we report parameter estimates, the maximum log-likelihood values and some model properties. Robust standard errors (based on outer product of gradients estimate) are in parantheses below the parameter estimates. We estimate six models. Each model has constant or time-varying volatility (which is either two components or one), and Normal or IG innovations.
### Table 1.3: Maximum Likelihood Estimation Results for Risk Neutral Distribution

#### Panel A. Parameter Estimates

<table>
<thead>
<tr>
<th>Model Type</th>
<th>Parameter Estimates</th>
<th>Maximum Log-Likelihood Values</th>
<th>Model Properties and Likelihoods</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gaussian Models</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Homoskedastic</td>
<td>( \hat{\mu} = -0.50 )</td>
<td>( \hat{\sigma} = 1.272E-04 ) (2.34E-7)</td>
<td></td>
</tr>
<tr>
<td>GARCH(1,1)</td>
<td>( \hat{\mu} = -0.50 )</td>
<td>( \hat{\sigma} = -1.260E-06 ) (6.99E-9)</td>
<td>( \hat{\beta}_1 = 0.823 ) (1.17E-3) ( \hat{\alpha}_1 = 2.931E-06 ) (8.76E-9) ( \hat{\gamma}_1 = 241.23 ) (8.78E-1)</td>
</tr>
<tr>
<td>GARCH(C)</td>
<td>( \hat{\mu} = -0.50 )</td>
<td>( \hat{\sigma} = 2.877E-07 ) (1.03E-8)</td>
<td>( \hat{\rho}_1 = 0.981 ) (1.95E-4) ( \hat{\gamma}_1 = 2096.18 ) (2.27E-8) ( \hat{\gamma}_2 = 0.9995 ) (8.12E+1)</td>
</tr>
<tr>
<td><strong>IG Models</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Homoskedastic</td>
<td>( \hat{\mu} = -0.48 )</td>
<td>( \hat{\sigma} = 1.593E-04 ) (2.67E-7)</td>
<td></td>
</tr>
<tr>
<td>IG-GARCH(1,1)</td>
<td>( \hat{\mu} = -0.50 )</td>
<td>( \hat{\sigma} = -1.956E-06 ) (2.68E-8)</td>
<td>( \hat{\beta}_1 = -2.50 ) (9.41E-3) ( \hat{\alpha}_1 = 4.931E+05 ) (5.32E+3) ( \hat{\gamma}_1 = 5.841E-06 ) (3.46E-8)</td>
</tr>
<tr>
<td>IG-GARCH(C)</td>
<td>( \hat{\mu} = -0.50 )</td>
<td>( \hat{\sigma} = 3.068E-07 ) (9.56E-9)</td>
<td>( \hat{\rho}_1 = 0.984 ) (1.51E-4) ( \hat{\gamma}_1 = 3.017E-06 ) (1.50E-5) ( \hat{\gamma}_2 = 0.9996 ) (3.15E-5)</td>
</tr>
</tbody>
</table>

#### Panel B: Model Properties and Likelihoods

<table>
<thead>
<tr>
<th>Model Type</th>
<th>Return Mean</th>
<th>Annualized Volatility</th>
<th>Volatility Persistence</th>
<th>Uncond. Skewness</th>
<th>Uncond. Kurtosis</th>
<th>Log Likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaussian Models</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Homoskedastic</td>
<td>1.69%</td>
<td>17.90%</td>
<td>0.000</td>
<td>3.000</td>
<td>32,632.1</td>
<td></td>
</tr>
<tr>
<td>GARCH(1,1)</td>
<td>0.21%</td>
<td>24.85%</td>
<td>0.993182</td>
<td>-0.017</td>
<td>5.113</td>
<td>53,971.2</td>
</tr>
<tr>
<td>GARCH(C)</td>
<td>-4.09%</td>
<td>38.44%</td>
<td>0.999990</td>
<td>-0.018</td>
<td>5.485</td>
<td>59,124.7</td>
</tr>
<tr>
<td>IG Models</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Homoskedastic</td>
<td>1.35%</td>
<td>20.04%</td>
<td>-0.101</td>
<td>3.699</td>
<td>36,945.6</td>
<td></td>
</tr>
<tr>
<td>IG-GARCH(1,1)</td>
<td>0.12%</td>
<td>25.24%</td>
<td>0.999322</td>
<td>-0.592</td>
<td>5.379</td>
<td>55,824.0</td>
</tr>
<tr>
<td>IG-GARCH(C)</td>
<td>-7.10%</td>
<td>45.61%</td>
<td>0.999994</td>
<td>-0.147</td>
<td>5.732</td>
<td>60,699.8</td>
</tr>
</tbody>
</table>

Parameter estimates are obtained by an MLE estimation on options from January 10, 1996 through December 26, 2012. Data description can be found in Table 1. For each model we report parameter estimates, the maximum log-likelihood values and some model properties. Robust standard errors (based on outer product of gradients estimate) are in parentheses below the parameter estimates. We estimate six models using only options data. Each model has constant or time-varying volatility (which is either two components or one), and Normal or IG innovations.
Table 1.4: Sequential Maximum Likelihood Estimation

Panel A. Risk-neutral parameters based on Table 2 and \( \xi \) estimates

<table>
<thead>
<tr>
<th>Gaussian Models</th>
<th>( \mu^* )</th>
<th>( \omega^* )</th>
<th>( \beta_1^* )</th>
<th>( \alpha_1^* )</th>
<th>( \gamma_1^* )</th>
<th>( \xi )</th>
<th>( \eta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARCH(1,1)</td>
<td>-0.50</td>
<td>-1.641E-06</td>
<td>0.900</td>
<td>5.193E-06</td>
<td>125.89</td>
<td>19791.7</td>
<td>1.1749</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(8.13E+1)</td>
<td></td>
</tr>
<tr>
<td>GARCH(C)</td>
<td>-0.50</td>
<td>2.445E-06</td>
<td>0.708</td>
<td>1.369E-06</td>
<td>725.92</td>
<td>21122.1</td>
<td>1.1930</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(8.17E+1)</td>
<td></td>
</tr>
<tr>
<td>IG Models</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IG-GARCH(1,1)</td>
<td>-0.50</td>
<td>-1.768E-06</td>
<td>-21.82</td>
<td>2.205E+07</td>
<td>5.854E-06</td>
<td>-6.886E-04</td>
<td>24225.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(9.05E+1)</td>
<td></td>
</tr>
<tr>
<td>IG-GARCH(C)</td>
<td>-0.50</td>
<td>2.415E-06</td>
<td>0.745</td>
<td>1.033E+06</td>
<td>9.682E-07</td>
<td>4.911E+07</td>
<td>4.660E+06</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(9.55E+1)</td>
<td></td>
</tr>
</tbody>
</table>

Panel B. Model Properties and Likelihoods

<table>
<thead>
<tr>
<th>Gaussian Models</th>
<th>Return Mean</th>
<th>Annualized Volatility</th>
<th>Volatility Persistence</th>
<th>Uncond. Skewness</th>
<th>Uncond. Kurtosis</th>
<th>Log Likelihood</th>
<th>LL Increase</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARCH(1,1)</td>
<td>0.84%</td>
<td>22.18%</td>
<td>0.981805</td>
<td>-0.013</td>
<td>4.745</td>
<td>48,065.9</td>
<td>6,643.5</td>
</tr>
<tr>
<td>GARCH(C)</td>
<td>0.67%</td>
<td>22.94%</td>
<td>0.996582</td>
<td>-0.017</td>
<td>5.190</td>
<td>51,034.7</td>
<td>9,336.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>IG Models</th>
<th>Return Mean</th>
<th>Annualized Volatility</th>
<th>Volatility Persistence</th>
<th>Uncond. Skewness</th>
<th>Uncond. Kurtosis</th>
<th>Log Likelihood</th>
<th>LL Increase</th>
</tr>
</thead>
<tbody>
<tr>
<td>IG-GARCH(1,1)</td>
<td>0.69%</td>
<td>22.84%</td>
<td>0.984584</td>
<td>-0.202</td>
<td>4.799</td>
<td>50,295.6</td>
<td>9,548.3</td>
</tr>
<tr>
<td>IG-GARCH(C)</td>
<td>0.55%</td>
<td>23.44%</td>
<td>0.997176</td>
<td>-0.154</td>
<td>5.238</td>
<td>52,280.0</td>
<td>11,180.2</td>
</tr>
</tbody>
</table>

Parameter estimates are obtained by an sequential MLE estimation on options from January 10, 1990 through December 26, 2012. Data description can be found in Table 1. For each model we report parameter estimates, the maximum log-likelihood values and some model properties. Robust standard errors (based on outer product of gradients estimate) are in parentheses below the parameter estimates. We estimate preference parameter \( \xi \) of four models using the returns-based parameters reported in Table 2 by applying the transformations (from physical to risk-neutral measure) mentioned in the appendix. Each model has constant or time-varying volatility (which is either two components or one), and Normal or IG innovations. The last column in Panel B reports the increase in log-likelihood going from a linear (\( \xi=0 \)) to nonlinear pricing kernel.
Table 1.5: Implied Volatility RMSE and Bias by Moneyness

Panel A. IV RMSE (Bias) by Moneyness for Models Fitted to Returns Only

<table>
<thead>
<tr>
<th>Model</th>
<th>F/X ≤ .80</th>
<th>.80 &lt; F/X ≤ .90</th>
<th>.90 &lt; F/X ≤ 1.00</th>
<th>1.00 &lt; F/X ≤ 1.10</th>
<th>1.10 &lt; F/X ≤ 1.20</th>
<th>F/X&gt;1.20</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARCH(1,1)</td>
<td>4.2651 (3.4014)</td>
<td>4.0436 (2.3690)</td>
<td>4.4883 (2.1799)</td>
<td>5.4407 (3.4356)</td>
<td>6.3429 (5.0015)</td>
<td>6.9426 (5.8150)</td>
<td>5.3289 (4.3274)</td>
</tr>
<tr>
<td>GARCH(C)</td>
<td>3.8684 (3.0676)</td>
<td>3.5813 (2.2484)</td>
<td>4.0457 (2.2724)</td>
<td>5.1913 (3.5101)</td>
<td>6.2316 (5.0933)</td>
<td>6.8989 (5.9136)</td>
<td>5.0694 (3.4766)</td>
</tr>
<tr>
<td>IG-GARCH(1,1)</td>
<td>4.2076 (3.4509)</td>
<td>3.8985 (2.6056)</td>
<td>4.3027 (2.4838)</td>
<td>5.3553 (3.5878)</td>
<td>6.2616 (5.0382)</td>
<td>6.8554 (5.8009)</td>
<td>5.2161 (3.5962)</td>
</tr>
<tr>
<td>IG-GARCH(C)</td>
<td>3.8121 (3.1466)</td>
<td>3.5629 (2.4280)</td>
<td>4.0048 (2.4566)</td>
<td>5.1453 (3.5727)</td>
<td>6.1570 (5.0590)</td>
<td>6.8110 (5.8486)</td>
<td>5.0179 (3.5610)</td>
</tr>
</tbody>
</table>

Panel B. IV RMSE (Bias) by Moneyness for Models Fitted to Options Only

<table>
<thead>
<tr>
<th>Model</th>
<th>F/X ≤ .80</th>
<th>.80 &lt; F/X ≤ .90</th>
<th>.90 &lt; F/X ≤ 1.00</th>
<th>1.00 &lt; F/X ≤ 1.10</th>
<th>1.10 &lt; F/X ≤ 1.20</th>
<th>F/X&gt;1.20</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARCH(1,1)</td>
<td>3.8553 (1.3958)</td>
<td>3.4887 (0.3345)</td>
<td>3.4791 (-0.1077)</td>
<td>3.5988 (0.4029)</td>
<td>4.1294 (1.9483)</td>
<td>4.7093 (2.9076)</td>
<td>3.7663 (0.7732)</td>
</tr>
<tr>
<td>GARCH(C)</td>
<td>3.4674 (2.2275)</td>
<td>2.9646 (1.3057)</td>
<td>2.7904 (0.1487)</td>
<td>3.1088 (-0.1795)</td>
<td>3.4522 (0.7840)</td>
<td>3.9121 (1.6311)</td>
<td>3.1545 (0.5035)</td>
</tr>
<tr>
<td>IG-GARCH(1,1)</td>
<td>3.8871 (1.0528)</td>
<td>3.3177 (0.1329)</td>
<td>3.0903 (-0.1521)</td>
<td>3.3615 (0.3737)</td>
<td>3.9667 (1.8961)</td>
<td>4.5821 (2.8858)</td>
<td>3.5336 (0.7051)</td>
</tr>
<tr>
<td>IG-GARCH(C)</td>
<td>3.4030 (1.3913)</td>
<td>2.8663 (0.6174)</td>
<td>2.5614 (-0.0955)</td>
<td>2.8312 (0.0758)</td>
<td>3.3569 (1.2359)</td>
<td>3.9144 (1.2352)</td>
<td>2.9875 (0.5460)</td>
</tr>
</tbody>
</table>

Panel C. IV RMSE (Bias) by Moneyness for Models Fitted to Options Sequentially

<table>
<thead>
<tr>
<th>Model</th>
<th>F/X ≤ .80</th>
<th>.80 &lt; F/X ≤ .90</th>
<th>.90 &lt; F/X ≤ 1.00</th>
<th>1.00 &lt; F/X ≤ 1.10</th>
<th>1.10 &lt; F/X ≤ 1.20</th>
<th>F/X&gt;1.20</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARCH(1,1)</td>
<td>3.7330 (1.7271)</td>
<td>4.3861 (-0.1290)</td>
<td>4.3634 (-0.4802)</td>
<td>4.3408 (0.8915)</td>
<td>5.0734 (2.8571)</td>
<td>5.7419 (4.0494)</td>
<td>4.6155 (1.0174)</td>
</tr>
<tr>
<td>GARCH(C)</td>
<td>3.5061 (0.9489)</td>
<td>3.9614 (-0.5154)</td>
<td>3.6967 (-0.5480)</td>
<td>3.9011 (0.7551)</td>
<td>4.7363 (2.7687)</td>
<td>5.5014 (4.0094)</td>
<td>4.1672 (0.8692)</td>
</tr>
<tr>
<td>IG-GARCH(1,1)</td>
<td>3.5286 (1.6067)</td>
<td>4.0133 (-0.1417)</td>
<td>3.8638 (-0.4725)</td>
<td>4.0715 (0.7085)</td>
<td>4.8059 (2.5988)</td>
<td>5.4717 (3.7719)</td>
<td>4.2747 (0.8913)</td>
</tr>
<tr>
<td>IG-GARCH(C)</td>
<td>3.4103 (0.9890)</td>
<td>3.7561 (-0.4212)</td>
<td>3.4667 (-0.4919)</td>
<td>3.7688 (0.6511)</td>
<td>4.5792 (2.5756)</td>
<td>5.3209 (3.7905)</td>
<td>3.9924 (0.8146)</td>
</tr>
</tbody>
</table>

We report implied volatility (IV) RMSE (values before parentheses) and bias (values inside parentheses) in percent by moneyness. The bias is defined as market IV less model IV. Panel A uses the parameter estimates from the return-based estimation in Table 2, Panel B uses the options-based estimates in Table 3, and Panel C uses the sequential estimates in Table 4.
### Table 1.6: Implied Volatility RMSE and Bias by Maturity

#### Panel A. IV RMSE (Bias) by Maturity for Models Fitted to Returns Only

<table>
<thead>
<tr>
<th>Model</th>
<th>DTM ≤ 30</th>
<th>30 &lt; DTM ≤ 60</th>
<th>60 &lt; DTM ≤ 90</th>
<th>90 &lt; DTM ≤ 120</th>
<th>120 &lt; DTM ≤ 180</th>
<th>DTM &gt; 180</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARCH(1,1)</td>
<td>4.6495 (2.9547)</td>
<td>5.2143 (3.3967)</td>
<td>5.2047 (3.3524)</td>
<td>5.6930 (3.9329)</td>
<td>5.3252 (3.5045)</td>
<td>5.5520 (3.4450)</td>
<td>5.3289 (3.4274)</td>
</tr>
<tr>
<td>GARCH(C)</td>
<td>4.4964 (3.0103)</td>
<td>5.0258 (3.4418)</td>
<td>4.9616 (3.3926)</td>
<td>5.3320 (3.8838)</td>
<td>5.0543 (3.5510)</td>
<td>5.2453 (3.5312)</td>
<td>5.0694 (3.4766)</td>
</tr>
<tr>
<td>IG-GARCH(1,1)</td>
<td>4.5423 (3.0684)</td>
<td>5.0795 (3.5273)</td>
<td>5.0601 (3.5077)</td>
<td>5.4981 (4.0699)</td>
<td>5.2431 (3.6863)</td>
<td>5.4791 (3.6685)</td>
<td>5.2161 (3.5962)</td>
</tr>
<tr>
<td>IG-GARCH(C)</td>
<td>4.4689 (3.0707)</td>
<td>4.9698 (3.5111)</td>
<td>4.8928 (3.4777)</td>
<td>5.2497 (3.9643)</td>
<td>5.0139 (3.6417)</td>
<td>5.2034 (3.6316)</td>
<td>5.0179 (3.5610)</td>
</tr>
</tbody>
</table>

#### Panel B. IV RMSE (Bias) by Maturity for Models Fitted to Options Only

<table>
<thead>
<tr>
<th>Model</th>
<th>DTM ≤ 30</th>
<th>30 &lt; DTM ≤ 60</th>
<th>60 &lt; DTM ≤ 90</th>
<th>90 &lt; DTM ≤ 120</th>
<th>120 &lt; DTM ≤ 180</th>
<th>DTM &gt; 180</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARCH(1,1)</td>
<td>4.3171 (1.3157)</td>
<td>4.2394 (1.3154)</td>
<td>3.7668 (0.8594)</td>
<td>3.6197 (1.0708)</td>
<td>3.4035 (0.6858)</td>
<td>3.4165 (0.1294)</td>
<td>3.7663 (0.7732)</td>
</tr>
<tr>
<td>GARCH(C)</td>
<td>4.0301 (0.8330)</td>
<td>3.7421 (0.9144)</td>
<td>3.2046 (0.5466)</td>
<td>2.9483 (0.6318)</td>
<td>2.7796 (0.4902)</td>
<td>2.5290 (0.0589)</td>
<td>3.1545 (0.5035)</td>
</tr>
<tr>
<td>IG-GARCH(1,1)</td>
<td>4.0868 (1.3671)</td>
<td>3.9773 (1.3381)</td>
<td>3.4912 (0.8563)</td>
<td>3.2858 (1.0133)</td>
<td>3.2160 (0.5547)</td>
<td>3.2401 (-0.0475)</td>
<td>3.5336 (0.7051)</td>
</tr>
<tr>
<td>IG-GARCH(C)</td>
<td>3.8771 (1.1875)</td>
<td>3.6164 (1.2356)</td>
<td>3.0171 (0.7659)</td>
<td>2.7624 (0.6886)</td>
<td>2.5303 (0.4240)</td>
<td>2.3618 (-0.2367)</td>
<td>2.9875 (0.5460)</td>
</tr>
</tbody>
</table>

#### Panel C. IV RMSE (Bias) by Maturity for Models Fitted to Options Sequentially

<table>
<thead>
<tr>
<th>Model</th>
<th>DTM ≤ 30</th>
<th>30 &lt; DTM ≤ 60</th>
<th>60 &lt; DTM ≤ 90</th>
<th>90 &lt; DTM ≤ 120</th>
<th>120 &lt; DTM ≤ 180</th>
<th>DTM &gt; 180</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARCH(1,1)</td>
<td>4.3879 (1.7672)</td>
<td>4.6955 (1.7208)</td>
<td>4.5165 (1.1667)</td>
<td>4.7743 (1.5080)</td>
<td>4.4305 (0.7540)</td>
<td>4.7115 (0.1848)</td>
<td>4.6155 (1.0174)</td>
</tr>
<tr>
<td>GARCH(C)</td>
<td>4.0062 (1.7594)</td>
<td>4.2872 (1.6855)</td>
<td>4.0786 (1.0797)</td>
<td>4.1949 (1.3008)</td>
<td>3.9768 (0.5875)</td>
<td>4.2529 (-1.007)</td>
<td>4.1672 (0.8692)</td>
</tr>
<tr>
<td>IG-GARCH(1,1)</td>
<td>4.0903 (1.7646)</td>
<td>4.3269 (1.6860)</td>
<td>4.1321 (1.0911)</td>
<td>4.2528 (1.3719)</td>
<td>4.1266 (0.6011)</td>
<td>4.4375 (-0.0544)</td>
<td>4.2747 (0.8913)</td>
</tr>
<tr>
<td>IG-GARCH(C)</td>
<td>3.9210 (1.7758)</td>
<td>4.1112 (1.6930)</td>
<td>3.8617 (1.0860)</td>
<td>3.9197 (1.2588)</td>
<td>3.8005 (0.5184)</td>
<td>4.1033 (-0.2402)</td>
<td>3.9924 (0.8146)</td>
</tr>
</tbody>
</table>

We report implied volatility (IV) RMSE (values before parentheses) and bias (values inside parentheses) in percent by maturity. The bias is defined as market IV less model IV. Panel A uses the parameter estimates from the return-based estimation in Table 2, Panel B uses the options-based estimates in Table 3, and Panel C uses the sequential estimates in Table 4.
Appendix A: Martingale Restrictions

A.1 Restrictions Implied by the Risk-free Asset

We first impose on the pricing kernel that the risk-free bond price is a martingale under the risk-neutral measure. We need

\[ E_t \left[ \frac{M(t + \Delta)}{M(t)} B_\tau(t + \Delta) \right] = B_\tau(t) \]

where \( B_\tau(t) \) is a bond with maturity \( \tau \) at time \( t \) and \( M(t + \Delta)/M(t) = (S(t + \Delta)/S(t))^\phi \exp(\delta(t + \Delta) + \xi h(t + 2\Delta)) \)

where \( \delta(t + \Delta) \equiv \delta_0 + \delta_1 h(t + \Delta) \). WLOG we assume that the risk-free rate is constant so that \( B_\tau(t + \Delta)/B_\tau(t) \equiv \exp(r_f) \). We can now write

\[ 1 = E_t \left[ \exp(\phi r(t + \Delta) + \delta_0 + \delta_1 h(t) + \xi h(t + 2\Delta) + r_f) \right]. \tag{1.45} \]

The martingale restriction in equation (1.45) implies that we need to impose the following parameter restrictions on the pricing kernel,

\[ \delta_0 = -(1 + \phi)r_f - \xi w + \frac{1}{2} \ln(1 - 2\xi a_1 \eta_4) \tag{1.46} \]

\[ \delta_1 = -\phi \mu - \xi b_1 - \eta^{-2} \left( 1 - \sqrt{(1 - 2\xi a_1 \eta_4)(1 - 2(\phi \eta + \xi c_1))} \right), \tag{1.47} \]

where we have used the following property of the IG distribution

\[ E_t[\exp(\alpha y(t + \Delta) + \beta/y(t + \Delta))] = \frac{1}{\sqrt{1 - 2\beta h(t + \Delta) - 2\eta^4}} \times \exp \left[ h(t + \Delta)/\eta^2 \left( 1 - \sqrt{(1 - 2\beta h(t + \Delta) - 2\eta^4)(1 - 2\alpha)} \right) \right]. \tag{1.48} \]

A.2 Restrictions Implied by the Risky Asset

We next impose on the pricing kernel that the risky stock is a martingale under the risk-neutral measure. We now need

\[ E_t \left[ \frac{M(t + \Delta)}{M(t)} S(t + \Delta) \right] = S(t) \]

We can thus write

\[ 1 = E_t \left[ \exp(\phi r(t + \Delta) + \delta_0 + \delta_1 h(t + \Delta) + \xi h(t + 2\Delta) + r(t + \Delta)) \right] \]
Taking logs this condition implies the following restriction on $\mu$,

$$\mu = \eta^{-2} \sqrt{(1 - 2\xi a_1 \eta a_4)} \left[ \sqrt{1 - 2(\eta + \phi + \xi c_1)} - \sqrt{1 - 2(\phi \eta + \xi c_1)} \right],$$

where we have used equation (1.48).
Appendix B: The Risk-Neutral Distribution

Consider the physical probability density function of the IG stock price

\[ f_{t-\Delta}(S(t)) = f_{t-\Delta}(y(t)) \left| \frac{\partial y(t)}{\partial S(t)} \right| \]

\[ = \frac{h(t)/\eta^3}{\sqrt{2\pi y(t)^3}S(t)} \exp \left( -\frac{1}{2} \left( \frac{\sqrt{y(t)} - h(t)/\eta^2}{\sqrt{y(t)}} \right)^2 \right) \]

(1.49)

To find the risk-neutral dynamic, we use the price kernel as follows

\[ f_{t-\Delta}^*(S(t)) = f_{t-\Delta}(S(t)) \exp(r_f)M(t)/M(t-\Delta), \]

(1.50)

where \( M(t-\Delta) \) is \((t-\Delta)\)-measurable.

Using the pricing kernel definition in (1.18) and the IG-GARCH(1,1) return dynamic, we can write

\[ f_{t-\Delta}^*(S(t)) = f_{t-\Delta}(S(t)) \exp\{[r_f + \delta_0 + \delta_1 h(t) + \phi \ln(S(t)/S(t-\Delta)) + \xi h(t + 2\Delta)]\} \]

\[ = \frac{h(t)/\eta^3 \sqrt{1 - 2\xi a\eta^4}}{\sqrt{2\pi y(t)^3}S(t)} \exp \left[ -\frac{1}{2} \left( \frac{\sqrt{1 - 2\phi \eta - 2\xi c} y(t) - h(t)/\eta^2}{\sqrt{y(t)}} \sqrt{1 - 2\xi a\eta^4} \right)^2 \right] \]

Substituting the physical distribution from equation (1.49) and rearranging terms yields

\[ f_{t-\Delta}^*(S(t)) = \frac{h(t) \sqrt{(1 - 2\xi a\eta^4)(1 - 2\phi \eta - 2\xi c)^{-3}(1 - 2\phi \eta - 2\xi c)^3/\eta^3}}{\sqrt{2\pi y(t)^3(1 - 2\phi \eta - 2\xi c)^3}S(t)} \]

\[ \times \exp \left[ -\frac{1}{2} \left( \frac{\sqrt{y(t)(1 - 2\phi \eta - 2\xi c)}}{\sqrt{1 - 2\phi \eta - 2\xi c}/\eta^2} \frac{h(t) \sqrt{(1 - 2\xi a\eta^4)(1 - 2\phi \eta - 2\xi c)^{-3}(1 - 2\phi \eta - 2\xi c)^3}/\eta^3}}{\sqrt{y(t)(1 - 2\phi \eta - 2\xi c)}} \right)^2 \right]. \]

This enables us to define the risk-neutral counterparts to \( y(t) \), \( h(t) \), and \( \eta \) by

\[ y^*(t) = y(t)(1 - 2\phi \eta - 2\xi c) = y(t)s_y, \]

\[ h^*(t) = h(t) \sqrt{(1 - 2\xi a\eta^4)(1 - 2\phi \eta - 2\xi c)^{-3}} = h(t)s_h, \]

\[ \eta^* = \eta/(1 - 2\phi \eta - 2\xi c) = \eta/s_y, \]
where we have used the definitions

\[ s_y = 1 - 2 \phi \eta - 2 \xi c \]
\[ s_h = \sqrt{1 - 2 \xi \alpha \eta^4 s_y^{-3/2}} \]

as in the text. Using these mappings yields the risk neutral density

\[ f^*_t(S(t)) = \frac{h^*(t)/|\eta^*|^3}{\sqrt{2\pi}(y^*(t))^3S(t)} \exp \left[ \frac{1}{2} \left( \frac{h^*(t)/(\eta^*)}{\sqrt{y^*(t)}} \right)^2 \right] \]

So that,

\[ f^*_t(y^*(t)) = f^*_t(S(t)) \left| \frac{\partial S(t)}{\partial y^*(t)} \right| = f^*_t(S(t)) |S(t) \times (-\eta^*)| \]

\[ = \frac{h^*(t)/(\eta^*)^2}{\sqrt{2\pi}(y^*(t))^3} \exp \left[ -\frac{1}{2} \left( \frac{h^*(t)/(\eta^*)}{\sqrt{y^*(t)}} \right)^2 \right]. \]

Therefore \( y^*(t) \) is distributed Inverse-Gaussian, and we can write,

\[ y^*(t) \sim IG \left( \frac{h^*(t)}{(\eta^*)^2} \right). \]

Using the physical return process and the above mappings we can write the risk-neutral return process as

\[ \ln(S(t + \Delta)) = \ln(S(t)) + r + \mu h^*(t + \Delta)/s_h + \eta y^*(t + \Delta)/s_y \]
\[ h^*(t + \Delta) = w s_h + b h^*(t) + c y^*(t)s_h/s_y + a s_y h^*(t)^2/y^*(t), \]

or equivalently

\[ \ln(S(t + \Delta)) = \ln(S(t)) + r + \mu^* h^*(t + \Delta) + \eta^* y^*(t + \Delta) \]
\[ h^*(t + \Delta) = w^* + b h^*(t) + c^* y^*(t) + a^* h^*(t)^2/y^*(t), \]

where we have used the parameter mapping in equation (1.21b) and (1.21c).
Appendix C: The Risk-Neutral Component Model

The component representation of the risk-neutral process (1.20) is given by

\[
\begin{align*}
\ln(S(t + \Delta)) &= \ln(S(t)) + r + \tilde{\mu}^* h(t + \Delta) + (\eta^* y^*(t + \Delta) - h^*(t + \Delta)/\eta^*), \\
h^*(t + \Delta) &= q^*(t + \Delta) + \rho_1^*(h^*(t) - q^*(t)) + \nu_q^*(t), \\
q^*(t + \Delta) &= \sigma^{*2} + \rho_2^*(q^*(t) - \sigma^{*2}) + \nu_q^*(t),
\end{align*}
\]

where

\[
q^*(t) = \frac{-\rho_1^* \tilde{w}^*}{(1 - \rho_1^*)(\rho_2^* - \rho_1^*)} + \frac{\rho_2^*}{\rho_2^* - \rho_1^*} h^*(t) + \frac{\tilde{b}_2^*}{\rho_2^* - \rho_1^*} h^*(t - \Delta) + \frac{1}{\rho_2^* - \rho_1^*} \nu_2^*(t - \Delta),
\]

\[
\tilde{\mu}^* = \mu^* + \eta^{* -1} = \mu/s_h + s_y/\eta^{* -1},
\]

\[
\sigma^{*2} = -\frac{\rho_2^*}{1 - \rho_2^*} \left( \tilde{w}^* \right),
\]

\[
\tilde{w}^* = w^* + a_1^* \eta^{*4} + a_2^* \eta^{*4} = s_h w + \frac{a_1^* \eta^{*4} + a_2^* \eta^{*4}}{s_h s_y^3},
\]

\[
\begin{align*}
\nu_h^*(t) &= c_h^* y^*(t) + a_h^* h^*(t)^2/y^*(t) - c_h^* h^*(t)/\eta^{*2} - a_h^* \eta^{*2} h^*(t) - a_h^* \eta^{*4}, \\
\nu_q^*(t) &= c_q^* y^*(t) + a_q^* h^*(t)^2/y^*(t) - c_q^* h^*(t)/\eta^{*2} - a_q^* \eta^{*2} h^*(t) - a_q^* \eta^{*4},
\end{align*}
\]

\[
\begin{align*}
a_h^* &= -\frac{\rho_1^*}{\rho_2^* - \rho_1^*} a_1^* h^* - \frac{1}{\rho_2^* - \rho_1^*} a_2^*, \\
c_h^* &= -\frac{\rho_1^*}{\rho_2^* - \rho_1^*} c_1^* h^* - \frac{1}{\rho_2^* - \rho_1^*} c_2^*, \\
a_q^* &= \frac{\rho_2^*}{\rho_2^* - \rho_1^*} a_1^* h^* + \frac{1}{\rho_2^* - \rho_1^*} a_2^*, \\
c_q^* &= \frac{\rho_2^*}{\rho_2^* - \rho_1^*} c_1^* h^* + \frac{1}{\rho_2^* - \rho_1^*} c_2^*, \\
\tilde{b}_1^* &= b_1 + s_h s_y c_1^*/\eta^{*2} + \frac{a_1^* \eta^{*2}}{s_h s_y},
\end{align*}
\]

and where \( \rho_1^* \) and \( \rho_2^* \) are the smaller and larger respective roots of the equation \( \rho^{*2} - \tilde{b}_1^* \rho - \tilde{b}_2^* = 0 \).
Appendix D: Conditional Moments

Consider the following basic definitions

\[ Var_t[h(t + 2\Delta)] \equiv E_t \left[ (h(t + 2\Delta) - E_t[h(t + 2\Delta)])^2 \right] \]

\[ Cov_t[R(t + \Delta), h(t + 2\Delta)] \equiv E_t [(R(t + \Delta) - E_t[R(t + \Delta)])(h(t + 2\Delta) - E_t[h(t + 2\Delta)])] \]

where \( R(t + \Delta) \equiv \ln S(t + \Delta) - \ln S(t) \).

In this section, we only focus on the derivation of conditional correlation, and conditional standard deviation of variance for IG-GARCH(C) model, since derivations for other models are similar.

Recall that the standardized conditional moments of an Inverse Gaussian random variable \( y(t + 1) \) are given by:

\[ E_t[y(t + \Delta)] = \delta(t + \Delta) \]

\[ Var_t[y(t + \Delta)] = \delta(t + \Delta) \]

\[ E_t[1/y(t + \Delta)] = 1/\delta(t + \Delta) + 1/\delta(t + \Delta)^2 \]

\[ Var_t[1/(t + \Delta)] = 1/\delta(t + \Delta)^3 + 2/\delta(t + \Delta)^4 \]

\[ Cov_t[y(t + \Delta), 1/y(t + \Delta)] = -1/\delta(t + \Delta), \]

where the degree of freedom is defined by

\[ \delta(t + \Delta) = h(t + \Delta)/\eta^2. \]

The variance process is defined as

\[ h(t + \Delta) = q(t + \Delta) + \rho_1 [h(t) - q(t)] + v_h(t) \]

\[ q(t + \Delta) = w_q + \rho_2 q(t) + v_q(t) \]

\[ v_h(t) = c_h [y(t) - \delta(t)] + a_h h(t)^2 \left[ 1/y(t) - 1/\delta(t) - 1/\delta(t)^2 \right] \]

\[ v_q(t) = c_q [y(t) - \delta(t)] + a_q h(t)^2 \left[ 1/y(t) - 1/\delta(t) - 1/\delta(t)^2 \right]. \]
Conditional variance of variance is given by

\[
Var_t[h(t + 2\Delta)] = (c_h + c_q)^2 h(t + \Delta)/\eta^2 - 2(a_h + a_q)(c_h + c_q)\eta^2 h(t) + (a_h + a_q)^2 \eta^6 h(t) + 2(a_h + a_q)^2 \eta^8.
\]

We thus can write

\[
Std_t[h(t + 2\Delta)] = \sqrt{2(a_h + a_q)^2 \eta^8 + [(c_h + c_q)/\eta - (a_h + a_q)\eta^3]^2 h(t + \Delta)}
\]

Consider now the innovation to returns

\[R(t + \Delta) - E_t[R(t + \Delta)] = \eta(y(t + \Delta) - \delta(t + \Delta))\]

We can then derive covariance and correlation

\[
Cov_t[R(t + \Delta), h(t + 2\Delta)] = (c_h + c_q)\eta Var_t[y(t + \Delta)] + (a_h + a_q)\eta h(t + \Delta)^2 Cov_t[y(t + \Delta), 1/y(t + \Delta)]
\]

\[= (c_h + c_q)/\eta h(t + \Delta) - (a_h + a_q)\eta^3 h(t + \Delta)\]

\[
Corr_t[R(t + \Delta), h(t + 2\Delta)] = \frac{Cov_t[R(t + \Delta), h(t + 2\Delta)]}{\sqrt{Var_t[R(t + \Delta)]Var_t[h(t + 2\Delta)]}}
\]

\[= \frac{[(c_h + c_q)/\eta - (a_h + a_q)\eta^3]\sqrt{h(t + \Delta)}}{\sqrt{2(a_h + a_q)^2 \eta^8 + [(c_h + c_q)/\eta - (a_h + a_q)\eta^3]^2 h(t + \Delta)}}\]
Chapter 2

The pricing of market jumps 
in the cross-section of equity options

2.1 Introduction

Sudden and large changes in stock prices cannot be explained by a normal shock (even with a time-varying variance) and studies have employed non-Gaussian distributions to capture these shocks\footnote{See e.g. Merton (1976), Naik and Lee (1990), Bates (1991), Bakshi, Cao and Chen (1997), Andersen, Bollerslev and Diebold (2003b)}. The seminal paper, Merton (1976) uses Compound Poisson jumps, however states that these jumps are likely to be non-systematic and could be diversified away in the cross-section of stocks. As a counter evidence, Ang and Chen (2002) documents that large negative returns coincide with increased cross-sectional stock return correlations. Furthermore, option pricing literature provides empirical evidence that jumps are systematic\footnote{See e.g. Bates (2000), Pan (2002), Broadie, Chernov and Johannes (2007)}. Recent studies\footnote{See e.g. Andersen, Bollerslev and Diebold (2007), and Barndorff-Nielsen and Shephard (2006)} show that high-frequency intraday data exhibit jumps in market returns.

If jumps are systematic, once a jump arrives into the market, stocks are expected to move with the market. Therefore, the jump sensitivity plays an important role in explaining individual stock returns and risk premiums. Todorov and Bollerslev (2010), and Bollerslev, Li and Todorov (2016) show that realized jump betas extracted from high-frequency intraday returns are superior to traditional betas.

However, the aforementioned papers are silent on stock options and the role of jump betas on stock options remains unknown. In this chapter, I develop a factor model to explain stock returns and option prices using market jumps.
in addition to normal shocks. The model allows for separate sensitivities of stock returns to market jumps and normal shocks and explains the role of normal and jump betas on the equity option cross-section for the first time in the literature.

Given the pricing kernel, I solve the model and show that higher betas imply higher expected returns due to the higher exposure to market risks. However, hedging against the market jumps gets more expensive at an increasing rate as the exposure increases. In addition, I derive an affine conditional moment-generating function (MGF) of future returns that reflects investors’ views about the future of the market and stocks. The MGF of stock returns is non-linear in market betas, but linear in the market and idiosyncratic variance components. The affine MGF enables me to derive a closed-form option pricing formula that greatly facilitates an empirical investigation of jump and normal market betas for equity options and returns.

In the empirical analysis, I use S&P 500 index as the market and the S&P 500 index constituent stock returns to look for evidence of jump betas. Moreover, in order to test the model-implied expectations, option contracts of 345 firms are used to fit the model.

The results show that modeling jump and normal market betas separately is important. First, the results of jump beta tests with model-free measures support that normal and jump betas are different for 65% of stocks at the 90% confidence level. Importantly, the jump beta in the model can explain the commonly observed time-varying CAPM-betas. Second, option pricing implications of the model indicate a difference in the effect of jump and normal betas as well. Higher jump betas increasingly raise option-implied volatilities as we go to deeper OTM for puts, whereas higher normal betas are unable to shift deep OTM puts but create an increase in ATM option prices.

Supporting evidence for a separate beta for market jumps is crucial because, first of all, ignoring jump beta can lead to time-varying CAPM-betas in stocks and these betas might change more rapidly as the contribution of jumps to the total variance increases. Furthermore, hedging against large negative market shocks using index options is more efficient if jump beta is introduced, because the separation of the two betas enables spanning of orthogonal risk factors. Moreover, jump betas directly link market volatility components to the total stock volatility, and common risk measures will be biased if jump betas are ignored. Finally, the market risk premium embedded in stock returns will not be estimated correctly if jump betas are ignored.

Although there is a rich literature that explores risk factors, attempts to understand if they are priced in the market, and whether the risks arising from normal shocks and jumps are priced differently there are few studies that focus on the pricing of these shocks in individual equities. Christoffersen et al. (2015) show there is a factor structure in equity options and develop a factor model to explain the pricing of market shocks in equity options. However,

---

4See e.g. Pan (2002), Bates (2008), Santa-Clara and Yan (2010)
their framework is limited to normal shocks and a higher beta implies both higher implied volatilities and steeper moneyness. Once jumps are introduced, I show that a higher jump beta causes steeper moneyness and a higher normal beta yields higher implied volatility.

To the best of my knowledge, the relative pricing of market risks in equity options as a function of factor sensitivities has not yet been explored. My paper introduces a new factor model to differentiate between stock return sensitivities to market jumps and normal shocks, and the model can generate separate effects of prices.

The remainder of the paper proceeds as follows. Section 2.2 describes the data sets and presents a preliminary analysis to support and justify the need for modeling a separate jump beta. Section 2.3 develops the model and discusses its implications. Section 2.4 shows the estimation methodology and section 2.5 presents and discusses the results. Section 2.6 concludes.

## 2.2 Data and preliminary analysis

A large cross-section of S&P 500 stock returns, in addition to stock options, market returns, and realized measures of variance data sets are employed in this paper. Section 2.2.1 describes these data sets, and section 2.2.2 presents an analysis of stock sensitivities to market shocks.

### 2.2.1 Description of data

As a proxy for the market returns, I use S&P 500 returns from January 3, 2000 to December 31, 2015. For the individual stock return data, I first use the S&P 500 index constituents dataset of Compustat to get the firms listed in S&P 500. Next, I obtain the ex-dividend returns data from CRSP Daily Stock file from January 3, 2000 to December 31, 2015. I select all 509 stocks with continuous trading (4025 trading days) during the date range of the market. To proxy for the risk-free interest rate, I use the 3-Month treasury bill in the same range as stocks and market returns.

For the equity option sample, I choose all of the available equity options in OptionMetrics from January 3, 2000 to the end of dataset (August 26, 2015) that has at least 1000 option contracts traded throughout the sample. These contracts are out-of-the-money (OTM) with moneyness (strike price divided by price of the underlying stock) between 0.75 and 1.25, and maturity between 30 and 90 days. Overall, more than 1.6 million options contracts with 345 underlying stocks are used in the estimations.
Realized measures of quadratic variation are downloaded from publicly available Oxford-Man Institute Realized Library. I select the realized variance and bipower variation time-series of S&P 500 index, calculated using intraday 5-minute returns with 1-minute subsampling for the period from January 3, 2000 to December 31, 2015.

Table 2.1 presents summary properties of each data set. Negatively skewed and leptokurtic nature of market returns are confirmed in Table 2.1. Stock returns are more volatile, even the lowest volatility stocks have higher volatility than the market. Option contracts have high IVs. And realized measures confirm that there is an additional variation from the jumps, since the difference between the means of realized variance and bipower variation is not negligible, it accounts for a tenth of realized volatility on average.

### 2.2.2 Is jump beta equal to normal beta?

It is well-documented that market returns exhibit jumps. However, this is a challenge to CAPM-type of models of stock returns, because stocks’ exposure to the market jumps might be different from the exposure to normal shocks. To see the potential problem, assume that CAPM holds

$$r_{j,t} - r_f = \beta_j (r_{m,t} - r_f) + w_{j,t}. \quad (2.1)$$

where $w_{j,t}$ is a zero mean idiosyncratic shock of stock $j$. If excess market returns consist of two sources of risk (i.e. a normal, $z$ and a jump, $y$), then the risk premium (RP) of stock $j$ can be expressed to be proportional to the market risk premium:

$$E[r_{j,t} - r_f] = \beta_j E[r_{m,t} - r_f] \quad (2.2)$$

$$RP_j = \beta_j RP_m = \beta_j (RP_{m,z} + RP_{m,y}) \quad (2.3)$$

which implies that stock returns have same sensitivity to market jumps and normal shocks. Stock returns can be written as

$$r_{j,t+1} = r_f + \beta_j RP_m + \beta_j (z_{t+1} + \bar{y}_{t+1}) + w_{j,t+1} \quad (2.4)$$

where $z_{t+1}$ and $\bar{y}_{t+1}$ are both zero mean shocks (normal and de-meaned jump) and the market returns are,

$$r_{m,t+1} = r_f + RP_m + z_{t+1} + \bar{y}_{t+1}. \quad (2.5)$$
Eq. (2.4) shows that sensitivities to all types of market jumps are the same under CAPM assumptions. However, one can allow for stock return innovations to have separate loadings on normal shocks and jumps:

\[ r_{j,t+1} = r_f + \beta_{z,j} R_{m,z} + \beta_{y,j} R_{m,y} + \beta_{z,j} \bar{y}_{t+1} + \beta_{y,j} \bar{y}_{t+1} + w_{j,t+1}. \]  

(2.6)

Eq. (2.6) has two implications if we do not differentiate between the innovations and simply calculate beta as the co-variation between stock and market returns: (i) CAPM-betas move with jump-to-total shocks to market returns, and (ii) CAPM-betas move with the ratio of variance contribution of jumps to total variance.

**CAPM-betas vs. jumps**

To see how CAPM-betas are related with jumps, Eq. (2.4) and (2.6) can be put together since both express the same on the left-hand side.

\[ \beta_j = \beta_{z,j} + (\beta_{y,j} - \beta_{z,j}) \frac{R_{m,y} + \bar{y}_{t+1}}{r_{m,t+1} - r_f}. \]  

(2.7)

Eq. (2.4) is based on CAPM specification, whereas Eq. (2.6) allows for separate loadings (betas). When Eq. (2.6) is the correct specification, Eq. (2.4) calculates as \( \beta_j \) is a function of the contribution of jumps to total returns and Eq. (2.7) shows that the CAPM-betas might fluctuate. According to Eq. (2.7), negative jumps of greater magnitude (given total returns are also negative) can cause increases in \( \beta_j \) if \( \beta_{y,j} - \beta_{z,j} > 0 \). Moreover, the required risk premium for market jumps (\( R_{m,y} \)) is likely to increase after a jump and therefore it constitutes more of the excess returns (\( r_{m,t+1} - r_f \)), which can explain the continuing rise even after the jumps. When there is no jumps in the market and \( R_{m,y} \) declines, this leads to a fall in \( \hat{\beta}_j \), because this time market excess returns are mostly explained by normal shocks and the risk premium required by the market. The next section offers an alternative approach to tackle this question, and the resulting relation can be tested using widely accepted model-free measures.

**CAPM-betas vs. variance by jumps**

The second implication of the misspecification of betas is that CAPM-beta is a linear function of variance ratios if Eq. (2.6) is the correct specification,

\[ \beta_j = \frac{Cov(r_{m,t+1}, r_{j,t+1})}{Var(r_{m,t+1})} = \beta_{z,j} + (\beta_{y,j} - \beta_{z,j}) \frac{Var(y_{t+1})}{Var(r_{m,t+1})}. \]  

(2.8)
I check whether CAPM-betas move with variance ratios. Each quarter, I calculate yearly CAPM-betas as done in section 2.2.2 and the ratio of “variation by discontinuous price moves” to “continuous variation” using model-free measures\(^6\). Variation by jumps can be captured by the difference between realized variance (RV) and realized bipower variation (BV), whereas total variation is measured by realized variance. Daily realized variance \(RV_t\) and realized bipower variation \(BV_t\) from Jan 2, 2000 to Dec 31, 2015 are calculated using 5-minute intraday returns of the S&P 500 index with 1-minute subsampling. To match the periods of \(\beta_j\) calculation, I take moving averages of \(RV_t\) and \(BV_t\) over a year. Figure 2.1 plots standardized (i.e. zero mean, unit variance) market averages of CAPM-betas and variance ratios. The two series seem to move together. Indeed, they have a correlation of 35%, which encourages further investigation to check whether there is a relation. The availability of data for variance ratios enables me to test the implications for individual firms.

In order to test Eq. (2.8), I perform a regression of the CAPM-betas on the variance ratios from the beginning of 2000 to the end of 2015,

\[
[beta]_q = \alpha_0 + \alpha_1 [JumpVar-to-TotalVar]_q + u_q. \tag{2.9}
\]

where CAPM-betas and the variance ratios are calculated quarterly. Table 2.2 summarizes the results from the regressions.

Panel A in Table 2.2 shows the results for three tests: (i) estimate of \(\alpha_0\) is significant, (ii) estimate of \(\alpha_1\) significant, and (iii) estimates of \(\alpha_0\) and \(\alpha_1\) are significant. Panel B presents the regressions with the highest and lowest \(R^2\)s (adjusted \(R^2\)s) for illustrative purposes\(^7\).

According to the results, the intercept of the regression, \(\alpha_0\), is significant in most cases, which is not surprising, because the stocks are the constituents of the S&P 500 index, from which the variance ratio is calculated. Since jumps are not observed frequently, the remaining normal shocks represent the market alone. Besides, the estimate of \(\alpha_1\) is significant at 10% for 332 stocks, and 322 of these stocks have a significant \(\alpha_0\) as well, meaning that the jump beta is significantly different from the normal beta for 65% of S&P500 firms unconditionally, which is quite striking. This is a strong indication that normal and jump betas are not the same.

Next section models the market return dynamics with normal shocks and jumps, and incorporates the evidence in this section into stock returns with separate market shock sensitivities for normal shocks and jumps.

---

\(^6\)See Barndorff-Nielsen and Shephard (2004) for the variation by discontinuous price moves. See also Andersen et al. (2003a), Barndorff-Nielsen and Shephard (2002a), Barndorff-Nielsen and Shephard (2002b), and Barndorff-Nielsen and Shephard (2005) for more about realized variance.

\(^7\)The full set of results of the 509 regressions are available upon request.
2.3 Model

Similar to the definitions of market and stock returns in section 2.2 Eqs. (2.5) and (2.6), market return innovations consist of a normal shock and a jump, whereas stock returns have exposures to market shocks in addition to their idiosyncratic shocks,

\[ r_{m,t+1} - E_t[r_{m,t+1}] = z_{t+1} + \bar{y}_{t+1} \quad (2.10) \]
\[ r_{j,t+1} - E_t[r_{j,t+1}] = \beta_{z,j} z_{t+1} + \beta_{y,j} \bar{y}_{t+1} + w_{j,t+1} \quad (2.11) \]

where \( z_{t+1} \) is a zero-mean normal shock with variance \( h_{z,t} \), \( \bar{y}_{t+1} \) is a demeaned compound Poisson jump (normal shocks with a Poisson arrival) with variance \( h_{y,t} \), and idiosyncratic \( w_{j,t+1} \) is a zero-mean normal shock with variance \( h_{j,t} \). All the shocks are orthogonal to each other. Stock return sensitivities to normal shocks (jumps) in market returns are \( \beta_{z,j} \) (\( \beta_{y,j} \)). Appendix presents the generalized model in detail.

The following section builds upon the market and stock return specifications, and introduces the variance components and the pricing kernel. The definition of the pricing kernel is important because forward-looking expectations are based on the preferences (i.e. pricing kernel), from which option prices can be determined.

2.3.1 Dynamics of variance components

Variance of stock returns are composed of normal-variance (\( h_{z,t} \)) and jump-variance (\( h_{y,t} \)) of market returns, and variance of stock specific shocks (\( h_{j,t} \)).

Market variance dynamics

Variance dynamics of normal shocks has the following dynamics:

\[ h_{z,t+1} = \sigma^2_z + \rho_z (h_{z,t} - \sigma^2_z) + a_z \left( \frac{z_{t+1}^2}{h_{z,t}} - 1 - 2 c_z z_{t+1} \right) . \quad (2.12) \]

which is equal to Heston-Nandi type GARCH(1,1)

\[ h_{z,t+1} = w_z + b_z h_{z,t} + a_z \left( \frac{z_{t+1}}{\sqrt{h_{z,t}}} - c_z \sqrt{h_{z,t}} \right)^2 \]

for \( \rho_z = b_z + a_z c_z^2 \) and \( \sigma^2_z = (w_z + a_z)/(1 - \rho_z) \).
Jump-variance (proportional to the jump intensity) could be modeled in several forms which are nested by the most generalized specification: (1) constant jump-variance, (2) proportional to normal-variance, (3) affine in normal-variance, (4) auto-regressive, and (5) random. These are:

\[ \text{Model 1} \quad h_{y,t+1} = \sigma_y^2 \]  
\[ \text{Model 2} \quad h_{y,t+1} = kh_{z,t+1} \]  
\[ \text{Model 3} \quad h_{y,t+1} = \sigma_y^2 + k(h_{z,t+1} - \sigma_z^2) \]  
\[ \text{Model 4} \quad h_{y,t+1} = \sigma_y^2 + \rho_y(h_{y,t} - \sigma_y^2) + k(h_{z,t+1} - \sigma_z^2) \]  
\[ \text{Model 5} \quad h_{y,t+1} = \sigma_y^2 + \rho_y(h_{y,t} - \sigma_y^2) + k(h_{z,t+1} - \sigma_z^2) + \frac{a_y}{h_{y,t}} \left[ y_{t+1}^2 - (1 + \nu_x^2 h_{y,t}) - 2\nu_y(y_{t+1} - \nu_x h_{y,t}) \right] \]  

Normal market shocks are expected to be correlated with their variance, \( h_{z,t} \) and the process moves slowly with relatively small shocks mostly, which makes \( h_{z,t} \) persistent. However, right after a negative jump in the market, total variance suddenly spikes up and stays at high levels for a while. Allowing only normal shocks to feed the normal-variance helps filtering clustered jumps (given the jump intensity is a function of jumps as in Model 5), because increased variance is captured by jump-variance (i.e. jump intensity) and following immediate jumps become more likely. An example is the sharp moves (mostly jumps) in the market in the late 2008.

**Variance dynamics of idiosyncratic shocks**

The variance dynamics of the idiosyncratic shocks to the stock \( i \)'s return is defined in different ways with nested models, similar to the jump-variance:

\[ \text{Model 1} \quad h_{i,t+1} = \sigma_i^2 \]  
\[ \text{Model 2} \quad h_{i,t+1} = k_i h_{z,t+1} \]  
\[ \text{Model 3} \quad h_{i,t+1} = \sigma_i^2 + k_i(h_{z,t+1} - \sigma_z^2) \]  
\[ \text{Model 4} \quad h_{i,t+1} = \sigma_i^2 + \rho_i(h_{i,t} - \sigma_i^2) + k_i(h_{z,t+1} - \sigma_z^2) \]  
\[ \text{Model 5} \quad h_{i,t+1} = \sigma_i^2 + \rho_i(h_{i,t} - \sigma_i^2) + k_i(h_{z,t+1} - \sigma_z^2) + a_i \left( \frac{w_{i,t+1}^2}{h_{i,t}} - 1 - 2c_i w_{i,t+1} \right) \]  

Idiosyncratic shocks variance are defined in a very flexible for in Model 5 and all other models are special cases of it. An interesting feature of the specifications with parameter \( k_i \) is that it allows a common factor in idiosyncratic

\[^{8}\text{Model 5 becomes Heston-Nandi type GARCH(1,1) when k = 0, similar to 2.12}\]
Recent studies investigate the variance of idiosyncratic shocks and find that they have a common market component. For the sake of model parsimony, I define this common factor as the divergence of the normal-variance from its unconditional mean but this specification can easily be extended to include other potential candidate state variables. The choice of normal-variance divergence could be justified by the common increase in the idiosyncratic variances during turmoils in the market, which is captured by the normal-variance.

### 2.3.2 The pricing kernel

Preferences toward each source of risk are the key elements to understand how the derivative contracts are priced. This makes the pricing kernel (stochastic discount factor – SDF) very crucial for the model. The innovations in the log-pricing kernel (log-SDF) are defined as an affine function of innovations to returns,

\[
m_{t+1} - E_t[m_{t+1}] = \Lambda_z z_{t+1} + \Lambda_y \bar{y}_{t+1} + \Lambda_h(z_{t+1} - E_t[h_{z,t+1}]) + \Lambda_h(y_{t+1} - E_t[h_{y,t+1}]). \tag{2.23}
\]

Investors dislike negative shocks to market returns and if there is a negative shock, marginal utility (and thus, log-SDF) increases. This implies the price of risk is negative for \(z_{t+1}\) and \(\bar{y}_{t+1}\), whereas a positive shock to \(h_{z,t+1}\) and/or \(h_{y,t+1}\) (not correlated with \(z_{t+1}\) and \(y_{t+1}\)), decreases investors’ utility and increases their marginal utility and thus the price of risk is positive for the shocks to \(h_{z,t+1}\) and \(h_{y,t+1}\).

### 2.3.3 The risk premium

The model implies that risk premia for the market and stock returns are affine in the variance components,

\[
\ln E_t[\exp(r_{m,t+1} - r_f)] = \mu_z h_{z,t} + \mu_y h_{y,t} \tag{2.24}
\]

\[
\ln E_t[\exp(r_{j,t+1} - r_f)] = \mu_{z,j} h_{z,t} + \mu_{y,j} h_{y,t} \tag{2.25}
\]

In Eq. (2.24), \(\mu_z h_{z,t}\) is the premium for the normal shocks borne by the market, whereas \(\mu_y h_{y,t}\) is for the jumps. Similarly, the risk premium in stocks can be decomposed into premium for normal shocks (\(\mu_{z,j} h_{z,t}\)) and premium for jumps (\(\mu_{y,j} h_{y,t}\)) in Eq. (2.25). The relation between risk sensitivities of stocks (\(\beta_{z,j}\) and \(\beta_{y,j}\)) and the price of risk (\(\mu_{z,j}\) and \(\mu_{y,j}\)) is very crucial, because it tells us how the market factors are priced in the cross-section of equity returns as a function of stock sensitivities and overall preferences of the market. To shed light on this

---

9See Serban et al. (2008), Schürhoff and Ziegler (2011), and Herskovic et al. (2014).
matter, I solve the martingale conditions. For each asset return satisfies the following equation,

\[ E_t[\exp(m_{t+1} + r_{j,t+1})] = 1. \]  

(2.26)

The solution for normal shocks are the following:

\[
\begin{align*}
\mu_z &= \psi_N(1) + \psi_{z,h} \left( \Lambda_z - 2(\Lambda_{hz} + \Lambda_{hy}k)a_z c_z, (\Lambda_{hz} + \Lambda_{hy}k)a_z \right) \\
&\quad - \psi_{z,h} \left( 1 + \Lambda_z - 2(\Lambda_{hz} + \Lambda_{hy}k)a_z c_z, (\Lambda_{hz} + \Lambda_{hy}k)a_z \right) \\
\mu_{z,j} &= \psi_N(\beta_{z,j}) + \psi_{z,h} \left( \Lambda_z - 2(\Lambda_{hz} + \Lambda_{hy}k)a_z c_z, (\Lambda_{hz} + \Lambda_{hy}k)a_z \right) \\
&\quad - \psi_{z,h} \left( \beta_{z,j} + \Lambda_z - 2(\Lambda_{hz} + \Lambda_{hy}k)a_z c_z, (\Lambda_{hz} + \Lambda_{hy}k)a_z \right).
\end{align*}
\]

The relations in case of jumps are:

\[
\begin{align*}
\mu_y &= \psi_y(1) + \psi_{y,h} \left( \Lambda_y - 2\Lambda_{hy}a_y c_y, \Lambda_{hy}a_y \right) - \psi_{y,h} \left( 1 + \Lambda_y - 2\Lambda_{hy}a_y c_y, \Lambda_{hy}a_y \right) \\
&\mu_{y,j} = \psi_y(\beta_{y,j}) + \psi_{y,h} \left( \Lambda_y - 2\Lambda_{hy}a_y c_y, \Lambda_{hy}a_y \right) - \psi_{y,h} \left( \beta_{y,j} + \Lambda_y - 2\Lambda_{hy}a_y c_y, \Lambda_{hy}a_y \right)
\end{align*}
\]

where \( \psi_y(c) \equiv (e^{c\mu_x + c^2\sigma_x^2/2} - 1)/(\mu_x^2 + \sigma_x^2) \).

In order to illustrate the relation, Figure 2.2 plots sensitivities of stocks to market shocks (\( \beta_{z,j} \) and \( \beta_{y,j} \)) against the corresponding values of \( \mu_{z,j} \) and \( \mu_{y,j} \). The relation is approximately linear as \( \beta_{y,j} \) increases when it is positive. However, when the sensitivity is negative, the magnitude of the price of risk increases non-linearly at an increasing rate. This means that an asset that appreciates when the market crashes is more expensive than an asset that moves counter to the market when the market receives a negative normal shock. An example of such assets could be OTM put options, as they gain value when a negative jump arrives into the market. Since they provide a hedge against market crashes, they are expensive and their expected returns are low. Figure 2.2 explains this phenomenon in a relation with jump sensitivity.

Since an OTM put option is a hedge against the risk source, its jump beta (OTM put option’s jump beta) is negative, when the underlying asset has positive sensitivity. And therefore, as the jump beta of the underlying increases, OTM put option’s jump beta becomes more negative. An implication is that higher sensitivity to market jumps requires paying more to hedge and therefore OTM puts are expected to be more expensive for stocks with higher jump sensitivity.

\[^{10}\text{See the appendix for the derivation}\]
2.3.4 Option pricing

Defining an affine structure in dynamics gives closed-form option prices written on stock \( j \)\(^{11} \):

\[
\text{Call}_{j,t}(X,\tau) = E^Q_t\left[e^{-\tau r_f, t}(e^{p_{j,t}} - X)^+\right] = -\frac{Xe^{-\tau r_f, t}}{\pi} \int_0^\infty \text{Re} \left[ \frac{\exp\left(M_{j,0}(\tau) + \sum_{l \in \{z, y, j\}} M_{j,l}(\tau) h_{l,t} + ik(x - p_{j,t})\right)}{k^2 - ik} \right] dk.
\]

where moment-generating function (MGF) of future returns are affine in state variables,

\[
M_{j,t}(u,\tau) = \exp\left(M_{j,0}(\tau) + M_{j,z}(\tau) h_{z,t} + M_{j,y}(\tau) h_{y,t} + M_{j,j}(\tau) h_{j,t}\right).
\]

Appendix shows the derivation of the MGF, which is an application of the law of iterated expectations, and discusses the option pricing formula in detail for both market and stocks. Log-MGF is affine in variance components, and the coefficients’ evolution through time is given as a recursive relation.

The previous section states that one implication of the relation between the price of jump risk and return sensitivity is that higher jump beta requires higher hedging cost, which implies higher implied volatilities for OTM put options. Given the option pricing formula, Figure 2.3 plots option-implied volatilities for several values of betas using model parameters that could create a contrast between normal and jump betas, and could reflect the flexibility of the model to capture tail behaviours.

The top left plot shows the effect of the jump beta when the normal beta is fixed, whereas jump beta is kept the same in the top right plot in order to illustrate the effect of the normal beta. The bottom left plot allows betas to move only together, and the bottom right presents the change in implied volatilities when the model is HN-GARCH(1,1) so that it does not have any jump variance \((h_{y,t})\) component. The top left plot supports the implication of Figure 2.2 that OTM put options get more expensive if they are more sensitive to jumps in the market, because the IV slope from ATM to OTM options gets steeper when the underlying asset has more sensitivity to market jumps.

The channel that makes an OTM put option more expensive for a given level of moneyness is the following: If the underlying asset return is more sensitive to market jumps then an OTM put option for a given level of moneyness has higher jump sensitivity and therefore is more expensive.

The top right plot, however, indicates more increase in ATM and less in OTM as the normal beta increases while keeping the the jump beta the same. In other words, the stocks that move more in the same direction with the

\(^{11}\)Stock options are American and they have early exercise feature. However, Bakshi, Kapadia and Madan (2003) shows that early exercise premium on equity options are negligible, and one can use European option pricing equations.
normal shocks to the market are expected to have higher IV level. This is also the case for the bottom right plot, where the underlying model does not have a jump component.

How does the CAPM-beta affect option contracts? As the bottom left plot shows, CAPM-beta (where jump and normal betas are the same) cannot replicate the separate effects that jump and normal betas can. Instead, the CAPM-beta can only generate close-to-parallel shifts in the IV smirk.

To sum up, the model implication in the top plots indicates that, jump and normal beta act differently in the pricing of option contracts. A higher jump beta implies steeper IV slope, and a higher normal beta implies higher IV levels.

The next section discusses the estimation methods used in determining stocks’ jump and normal betas.

2.4 Estimating the betas

This section describes the estimation methodology to filter the normal shocks and jumps in the market, as well as the market variance components, establishes an inference over the model parameters of the market, and determine the estimates of model-implied jump and normal betas for each stock, together with the filtered idiosyncratic variances.

2.4.1 Estimating the model on market returns

While there are several methods for parameter estimation such as efficient method of moments\(^{12}\) and approximate maximum likelihood\(^{13}\) filtering latent state variables optimally has recently been investigated in the literature\(^{14}\)

Since market returns have two sources of randomness and each shock feed different state variable, traditional GARCH filter does not apply to the model introduced in the previous sections. A remedy to this issue is to use particle filter (PF). PF is a method to approximate the likelihood function and filter the unobserved state variables. I estimate the proposed models using market returns, following the PF in Ornthanalai (2014), which is a SIR (sampling importance re-sampling) PF that samples 8192 particles, computes importance weights and re-samples based on the probability weights\(^{15}\)

\(^{12}\)See e.g. Gallant and Tauchen (1996)  
\(^{13}\)See e.g. Bates (2006)  
\(^{14}\)See e.g. Johannes, Poisson and Stroud (2009) for a review  
\(^{15}\)See Appendix A in Ornthanalai (2014).
The log-likelihood function is the sum of the total of non-normalized weights, \( \tilde{\pi} \):

\[
\hat{L}(\theta_m|r_{m,1:t}) = \hat{L}(\theta_m|r_{m,1:t-1}) + \ln \sum_{i=1}^{M} \tilde{\pi}_i^{(i)},
\]

(2.27)

where \( \theta_m \) is the set of parameters of the model.

Estimation results on market returns are crucial, because the filtered variance components and shocks are used in the maximum likelihood estimation (MLE) on stock returns. Since any misspecification in this step would be carried onto the next step (plus, the resulting log-likelihood values are significantly higher than others), I employ the most general specification of the model (i.e. model 5).

### 2.4.2 Maximum likelihood estimation of stock return dynamics

Since the model assumes that market shocks are independent from the idiosyncratic equity shocks, estimation of the model can be separated into two parts

\[
p(r_{m,1:T}, r_{j,1:T}|\theta_m, \theta_j) = p(r_{m,1:T}|\theta_m) p(r_{j,1:T}|\theta_j).
\]

Given the filtered state variables, market parameter estimates and market returns, estimating the stock parameter is to find the set of stock parameters of equity return models, that maximizes the likelihood of stock returns,

\[
\hat{\theta}_j = \arg \max_{\theta_j} \hat{L}(\theta_j|r_{j,1:T}, \hat{\theta}_m, \hat{h}_z, \hat{h}_y, \hat{z}, \hat{y}).
\]

(2.28)

### 2.4.3 Challenges in the estimations

The above estimation has challenges due to the use of daily returns. Parameters related with the risk premiums are difficult to differentiate from zero, when the state variables are latent. For this reason, I ignore \( \mu_z \) and \( \mu_y \) in the market model, and \( \mu_{z,i} \)'s and \( \mu_{y,i} \)'s in the equity model by setting \( \lambda_z = \lambda_y = \lambda_{hz} = \lambda_{hy} = 0 \), where market returns are

\[
r_{m,t+1} = r_{f,t} + \lambda_z h_{z,t} + \lambda_y h_{y,t} + z_{t+1} + y_{t+1}
\]

\( \tilde{\pi}_i^{(i)} \) is probability of \( z \) for a given \( y_i^{(i)} \) that is randomly generated.
and stock returns are

\[ r_{i,t+1} = r_{f,t} + \lambda_{z,i} h_{z,t} + \lambda_{y,i} h_{y,t} + \beta_{z,i} z_{t+1} + \beta_{y,i} y_{t+1} + w_{i,t+1}. \]

Results of the estimations with and without the above restriction are not significantly different based on Likelihood Ratio (LR) test.\[17\]

### 2.4.4 Betas and option fit

The model can be tested on options by estimating the pricing kernel parameters. Using the estimates from return observations and the filtered state variables, option prices and the root mean squared error (RMSE) can be calculated for a given set of pricing kernel parameters.

For statistical inference, Black-Scholes implied volatility (IV) of an option with strike price \( K \) and maturity \( \tau \) is assumed to be measured with error as,

\[ \hat{IV}(K, \tau) = IV(K, \tau) + \sigma \varepsilon_i \]

where \( \varepsilon_i \) is a standard normal shock. Therefore, each set of parameters and filtered states gives a likelihood value, which enables us to compare the model with the single beta (jump and normal betas are same) alternative.\[18\] The option log-likelihood is

\[ \ln L_\theta \propto -\frac{1}{2} \sum_i (\ln \sigma_\varepsilon^2 + \varepsilon_i^2 / \sigma_\varepsilon^2). \]

### 2.5 Results

This section presents results of the estimations described in section 2.4. The first estimation is to find the parameters of market dynamics. This estimation is run under the five models introduced previously and GARCH(1,1) model as a base case without jumps. Second, stock return dynamics are estimated, given the results from the first set of estimations of market dynamics. And finally, pricing kernel parameters estimated using the individual equity option contracts under specification with common beta (normal beta is equal to jump beta) and separate betas.

\[17\]Results without the restrictions could be provided upon request.
\[18\]See Christoffersen et al. (2013) for the details of the estimation. I use the same likelihood function for options, though I invert Black-Scholes IVs from model-implied option prices using Newton method and calculate the error \( \varepsilon_i \) directly.
2.5.1 Market return dynamics

Estimating the market dynamics gives the filtered shocks as in Figure 2.4 and filtered states in Figure 2.5. Figure 2.4 indicates that market rarely receives jump shocks, and these are mostly negative. Once a jump arrives, it is more likely for another jump to arrive. This mechanism is captured in model 5 where a jump arrival increases the intensity of future jumps. Therefore, one can expect for model 5 to have the highest log-likelihood value, since it captures the self-exciting nature of jumps by construction, and clustered jumps are most likely in model 5. Figure 2.5 confirms this prediction by showing the annualized standard deviations (volatility) of normal shocks and jumps. Average value of jump variance \( h_{y,t} \) is the highest in model 5, and the average of jump variance contribution to total variance \( h_{y,t}/(h_{z,t} + h_{y,t}) \) is 10% in model 5, which is higher than other models.

While Panel A of table 2.3 shows the parameter estimates of the nested models, Panel B summarizes the filtered state variables and compares the models.

According to Panel A, Model 5 has the lowest unconditional variance of normal shocks and it also has the highest unconditional variance of jumps, as predicted. Besides, models with jump have \( k \) parameter estimates between 4.4% and 5.2%, which means that divergence of variance of normal shocks from their unconditional mean contributes to the jump variance. Thus, during the times of slowly rising market volatility (due to the increase in normal variance), probability of jump arrivals increases. In other words, jumps could also be a result of prolonged uncertainty in the market (which might be stemming from macroeconomic conditions), in addition to unexpected arrivals.

Panel B sheds light into the importance of modeling jump dynamics, since it shows the log-likelihood ratio (LR) test results of the models 1 - 4 in a comparison with the most generalized specification (i.e. model 5). At 5% significance models 1 - 4 are all different from model 5, whereas only GARCH(1,1) and model 1 are different at 1% significance level.

Contribution of jump variance to the total variance is between 5.13% to 9.74% and average jump variance (and its standard deviation) increases with the flexibility of the model. Similarly, variance of the normal shocks is lower for more generalized model specifications.

\[ \text{Reference: Alternative, one can use high-frequency intraday returns to capture jumps and state variables. While state variables look similar to model filtered results, jumps in returns show big difference. See e.g. the realized jumps in Tauchen and Zhou (2011) Figure 4. They are very frequently observed and not only large negative jumps but there are many small and positive ones.} \]
Are the jumps rare, negative and large in magnitude?

The mean jump size estimates are between $-3.1\%$ and $-2.2\%$, and the jump standard deviation parameter estimates are between $2.5\%$ and $3.1\%$. The parameter estimates related with jumps, therefore, imply that the jumps are rare, but large and negative, which can also be seen in Figure 2.4. Negativity is important, because this is the key element that makes the price of risk a convex function of beta if the exposure is negative. Therefore, as the exposure of the asset to the jumps increases, the amount that the investors are willing to pay to hedge against the jump increases more. Note that in the normal shock case this relation is linear and symmetric.

What is the underlying process for market returns and volatilities?

Results in Table 2.3 show that the improvement by model 5 from all other models is significant at 5%. First of all, all the jump models improve GARCH(1,1) to a large extent. This is a simple indication of the jumps’ existence. Second, the more generalized models do significantly better as LR test values suggest.

2.5.2 Equity return dynamics

In order to estimate individual equity return dynamics, I use the resulting state variables of market dynamics under model 5. Given the market variables, I estimate model 5 of equity return dynamics.

Panel A of Table 2.4 summarizes the results of 509 estimations by showing 5%, 50%, and 95% of the estimates of each parameter. There is a dispersion in the values of betas. On average the jump beta is close to normal beta. There is a high correlation of 0.57 between the two betas. Among the other parameters, it is worth to mention the persistence of equity volatility dynamic. All the values are very high and the average is around 98%. Additionally, a firm’s own (idiosyncratic) volatility has a very high unconditional value: the estimate of $\sigma_j$, is 0.0185 which is 75% higher than the market normal volatility.

Is there a relation between jump and normal betas? How are they related with other parameters?

Table 2.4 shows the correlations of the stock parameters in Panel B, and p-values of the relations in Panel C. Although the correlation between jump beta and normal beta are high and the relation is significant, there are differences in how other parameters interact with the betas. An example is the persistence parameter $\rho_j$. It has a very significant relations with normal and jump betas, but the correlation coefficient with jump beta is not as high as normal beta. Therefore, there are differences in how these two betas interact with other parameters. Particularly
for the persistence, a strong relation with normal beta is a sign that the model captures slowly evolving dynamics with normal beta.

**How important is the market jump beta?**

One can check whether defining a separate beta to represent jump sensitivities significantly improves the model. For this purpose, a restricted model with a common beta (jump beta is set to equal the normal beta) is also estimated and likelihood ratio tests are performed to see what fraction of firms would favor the jump beta. According to Panel C of Table 2.4, the null hypothesis of the model with a common beta is rejected for 270 stocks at 10% significance, and this corresponds to 53% of all stocks.

### 2.5.3 Evidence from options

One of the important implications of the model is that the return sensitivities to jumps and normal shocks impact option prices. This implication can be tested on individual equity options. Given the estimates of market and individual equity parameters, pricing kernel parameters are estimated for each set of equity options. Table 2.5 summarizes the results of 345 individual equity option estimations with a total number of 1,605,189 option contracts from January 3, 2000 to August 26, 2015.

Panel A shows the mean and standard deviations of pricing kernel parameter estimates for model with separate betas and common beta. The estimates are similar to each other (within a very small range compared to the standard deviations). Therefore, the model choice of setting normal and jump betas equal does not affect the pricing kernel parameters. The price of risk parameter of normal shocks $\Lambda_z$ is only $-3.54$, whereas $\Lambda_y$ is $-31.34$. This is striking because the preference of the market towards the source of these two shocks is different. This is not easy to justify with standard utilities, unless we assume, for example, recursive preferences of Epstein and Zin (1989).

In the long-run risk framework of Bansal and Yaron (2004), this can be justified with the following specification: consumption and cash flow growth has a normal shock only, but the variance of this shock is stochastic and jumps can affect the variance\[^{20}\] Therefore, returns of the asset (market portfolio) that give the consumption as a dividend, are exposed to the shocks to the variance of consumption.

Panel B summarizes RMSE and log-likelihood values from 345 option estimations. According to the results, the difference between the likelihood of model with separate betas and that with common (same) beta is very significant, though the resulting averages of the RMSEs are similar. It could be interesting to look at equity

\[^{20}\]See Drechsler and Yaron (2011)
option estimations individually and test the restricted model of common beta. Panel C shows the percentages of estimations where the restricted model with common beta is rejected at 1%, 5%, and 10% significance levels. Results are supporting the previous section’s LR test findings. Around 41% of the individual equity option estimations rejects the common beta at 1% significance level, 47% at 5%, and 50% at 10%.

Overall, estimation results of equity options support that sensitivity to market shocks might differ and a model with separate loadings on normal shocks and jumps can better explain the pricing of the individual equity option contracts.

2.6 Conclusion

In this paper, I propose a new factor model to explain equity returns using market factors. I define a separate stock return sensitivity for market jumps. An analysis based on model-free measures of quadratic variation is presented in the paper to support the need to model market jump sensitivities separately.

Given the model, I derive conditional moment-generating function of future returns which gives a closed-form solution for option prices. Option pricing implications of the model is that jump betas have an impact mostly on OTM puts, whereas normal beta affects the option price less as the put options of interest gets deeper OTM. An explanation for this pricing difference is provided via the price of risk for each shocks. Based on the pricing kernel, I solve the model to find the market price of risk embedded in equities. The resulting relations are interesting because normal shocks yield a linear relationship showing that more exposure implies higher price. However, an asset that moves opposite of the jumps requires a price of risk that increases at an increasing rate.

Empirically, I employ a particle filter to extract the variance components from return observations, and compare five nested models. Moreover, I use a very large stock return data set. Using more than 2 million return observations from 509 constituents of S&P 500, I estimate the individual equity parameters. And finally, with the help of 1.6 million equity option contracts from 345 underlying stocks, I estimate the pricing kernel parameters.

Results show that jumps are very important in explaining the market and equity returns dynamics. All the market model specifications with jumps performs significantly better than the model without jumps. The return sensitivities of the assets to market shocks differ. Both the model-free method in section 2 and the estimations on equity returns support the use of separate betas. And this is also reflected in the equity options.
2.7 Figures

Figure 2.1: Standardized CAPM-betas and variance ratios.

Notes: Horizontal axis is years from beginning of 2000 to end of 2015, vertical axis is the standardized values (where each variable in the figure has zero mean and unit variance). CAPM-betas are calculated at each quarter (63 trading days) using the most recent one-year long (252 trading days) observations. Plotted values are average CAPM-betas across all firms for each quarter, having an average of 1.0164. Variance ratio is the difference between one-year moving averages of RV and BV, divided by RV. Its mean is 0.1134 and its correlation with CAPM-beta averages is 0.3477.
Figure 2.2: $\beta_{z,j}, \beta_{y,j}$ vs. $\mu_{z,j}, \mu_{y,j}$.

Notes: Horizontal axis is the sensitivity to market shocks ($\beta_{z,j}$ for normal shocks, and $\beta_{y,j}$ for jumps), vertical axis is price of risk ($\mu_{z,j}$ for normal shocks, and $\mu_{y,j}$ for jumps). $\Lambda_z$ and $\Lambda_y$ values are chosen as 10, $\Lambda_{hz}$ and $\Lambda_{hy}$ are 500, and average jump size $\mu_x$ is -2%, standard deviation of jumps $\sigma_x$ is 2.5%.
Figure 2.3: Jump and diffusive $\beta$ vs. implied volatilities.

Notes: Horizontal axis is moneyness (strike price divided by current price of the underlying), vertical axis is implied volatilities. Top left plot fixes $\beta_z = 1$ and shows IVs for $\beta_y$s. Top right plot fixes $\beta_y = 1$ and shows IVs for $\beta_z$s. Bottom left plot shows IVs for $\beta_z = \beta_y$. Bottom right plot cancels the jump, it is based on a model with only normal shocks.
Figure 2.4: Filtered S&P 500 index return shocks

Notes: Horizontal axis is time in all graphs. Vertical axis is daily returns. Gray lines are normal shocks, black lines are jumps.
Figure 2.5: Filtered state variables: Volatility of normal shocks and jumps of S&P 500 index returns

Notes: Horizontal axis is time and vertical axis is annualized volatility in all graphs. Black lines plot the annualized volatility of normal shocks ($\sqrt{h_{z,t}} \times 252$), red lines are of jumps ($\sqrt{h_{y,t}} \times 252$)
2.8 Tables
Table 2.1: Descriptive statistics

Panel A: Market variables

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Number of Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>S&amp;P 500 Returns</td>
<td>0.0207</td>
<td>0.2011</td>
<td>-0.1848</td>
<td>11.0044</td>
<td>4025</td>
</tr>
<tr>
<td>Risk-free rate</td>
<td>0.0119</td>
<td>0.0009</td>
<td>0.8931</td>
<td>2.3765</td>
<td>4025</td>
</tr>
</tbody>
</table>

Panel B: Stock return percentiles

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>5%</td>
<td>-0.0854</td>
<td>0.2285</td>
<td>-2.1898</td>
<td>7.6156</td>
</tr>
<tr>
<td>50%</td>
<td>0.0541</td>
<td>0.3837</td>
<td>-0.1762</td>
<td>13.0041</td>
</tr>
<tr>
<td>95%</td>
<td>0.1905</td>
<td>0.6633</td>
<td>0.4185</td>
<td>57.8634</td>
</tr>
</tbody>
</table>

Total number of obs. : 2,048,725. Number of stocks : 509. Number of obs. per stock : 4025

Panel C: Stock option IV percentiles of the moments

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Number of Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>5%</td>
<td>0.2259</td>
<td>0.0820</td>
<td>0.9110</td>
<td>3.6405</td>
<td>1156</td>
</tr>
<tr>
<td>50%</td>
<td>0.3457</td>
<td>0.1196</td>
<td>1.6858</td>
<td>7.0492</td>
<td>3675</td>
</tr>
<tr>
<td>95%</td>
<td>0.5045</td>
<td>0.2159</td>
<td>2.6127</td>
<td>14.5053</td>
<td>10339</td>
</tr>
</tbody>
</table>

Total number of option contracts : 1,605,189. Number of underlying stocks : 345

Panel D: High-frequency realized measures of variation

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Number of Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Realized variance</td>
<td>1.0321</td>
<td>2.3790</td>
<td>15.3052</td>
<td>432.5740</td>
<td>3996</td>
</tr>
<tr>
<td>Bipower variance</td>
<td>0.9286</td>
<td>2.2309</td>
<td>15.0868</td>
<td>410.9506</td>
<td>3996</td>
</tr>
</tbody>
</table>

Notes: Panel A shows first four moments of S&P 500 returns and 3-month T-bill rate. Both time-series are from January 3, 2000 to December 31, 2015. Panel B summarizes 509 time-series of stock returns. For the number of observations and each return moment 5%, 50% and 95% percentiles are shown. Similarly, Panel C presents the option-implied volatilities of the option data by showing the values at 5%, 50% and 95% percentiles. Index Constituents from January 2000 to December 2015 are obtained from Compustat. Ex-dividend stock returns data are from CRSP Daily Stock file from January 3, 2000 to December 31, 2015. Option contracts are obtained from OptionMetrics (January 3, 2000 and August 26, 2015). Panel D shows the first four moments of realized variance and bipower variation. Data are calculated using intraday 5-minute returns with 1-minute subsampling for the period from January 3, 2000 to December 31, 2015, obtained from Oxford-Man Institute Realized Library.
Table 2.2: Regression results

Panel A: Summary statistics: Significant results

<table>
<thead>
<tr>
<th>Significance level</th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_0 \neq 0$</td>
<td>434</td>
<td>481</td>
<td>488</td>
</tr>
<tr>
<td>$\alpha_1 \neq 0$</td>
<td>158</td>
<td>277</td>
<td>332</td>
</tr>
<tr>
<td>$\alpha_0$ and $\alpha_1 \neq 0$</td>
<td>112</td>
<td>262</td>
<td>322</td>
</tr>
</tbody>
</table>

Number regressions: 509

Panel B: Results of regressions with highest and lowest Adj $R^2$

<table>
<thead>
<tr>
<th>Ticker</th>
<th>$\hat{\alpha}_0$ (t-stat)</th>
<th>$\hat{\alpha}_1$ (t-stat)</th>
<th>F-stat (p-val)</th>
<th>$R^2$ (Adj $R^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(highest 5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TEX</td>
<td>0.444 (1.676)</td>
<td>11.548 (5.360)</td>
<td>28.725 (0.000)</td>
<td>0.320 (0.309)</td>
</tr>
<tr>
<td>BCO</td>
<td>0.390 (2.822)</td>
<td>5.577 (4.950)</td>
<td>24.505 (0.000)</td>
<td>0.287 (0.275)</td>
</tr>
<tr>
<td>CSC</td>
<td>0.548 (4.394)</td>
<td>4.369 (4.302)</td>
<td>18.507 (0.000)</td>
<td>0.233 (0.220)</td>
</tr>
<tr>
<td>FMC</td>
<td>0.587 (4.811)</td>
<td>4.254 (4.286)</td>
<td>18.368 (0.000)</td>
<td>0.231 (0.219)</td>
</tr>
<tr>
<td>PPG</td>
<td>0.671 (6.801)</td>
<td>3.411 (4.249)</td>
<td>18.051 (0.000)</td>
<td>0.228 (0.216)</td>
</tr>
<tr>
<td>(lowest 5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WDC</td>
<td>1.471 (6.617)</td>
<td>-0.034 (-0.019)</td>
<td>0.000 (0.493)</td>
<td>0.000 (-0.016)</td>
</tr>
<tr>
<td>STR</td>
<td>0.789 (4.289)</td>
<td>-0.053 (-0.036)</td>
<td>0.001 (0.486)</td>
<td>0.000 (-0.016)</td>
</tr>
<tr>
<td>RHT</td>
<td>1.334 (4.470)</td>
<td>-0.100 (-0.041)</td>
<td>0.002 (0.484)</td>
<td>0.000 (-0.016)</td>
</tr>
<tr>
<td>AMD</td>
<td>1.700 (5.858)</td>
<td>0.106 (0.045)</td>
<td>0.002 (0.482)</td>
<td>0.000 (-0.016)</td>
</tr>
<tr>
<td>JBHT</td>
<td>1.120 (6.482)</td>
<td>0.066 (0.047)</td>
<td>0.002 (0.481)</td>
<td>0.000 (-0.016)</td>
</tr>
</tbody>
</table>

Notes: Regression is $[\beta_q]_q = \alpha_0 + \alpha_1[\text{JumpVar-to-TotalVar}]_q + u_q$. Panel A shows the number of regressions that have significant results for three tests: (i) estimate of $\alpha_0$ is significant, (ii) estimate of $\alpha_1$ significant, and (iii) estimates of $\alpha_0$ and $\alpha_1$ are significant. Panel B presents the regressions with the highest and lowest $R^2$s (adjusted) for illustrative purposes.
Table 2.3: Estimation results of market return and volatility dynamics

**Panel A: Parameter Estimates**

<table>
<thead>
<tr>
<th></th>
<th>$\sigma_z \times 100$</th>
<th>$\rho_z$</th>
<th>$\alpha_z \times 10^6$</th>
<th>$\epsilon_z$</th>
<th>$k$</th>
<th>$\sigma_y \times 100$</th>
<th>$\rho_y$</th>
<th>$\alpha_y \times 10^6$</th>
<th>$\epsilon_y$</th>
<th>$\mu_y$</th>
<th>$\sigma_y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARCH(1,1)</td>
<td>1.1736</td>
<td>0.9792</td>
<td>3.8884</td>
<td>205.15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1</td>
<td>1.0762</td>
<td>0.9859</td>
<td>2.9030</td>
<td>264.34</td>
<td>0.1983</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 2</td>
<td>1.0871</td>
<td>0.9846</td>
<td>3.1171</td>
<td>235.92</td>
<td>0.0519</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 3</td>
<td>1.0895</td>
<td>0.9856</td>
<td>3.1978</td>
<td>235.19</td>
<td>0.0438</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 4</td>
<td>1.0801</td>
<td>0.9856</td>
<td>3.1447</td>
<td>239.58</td>
<td>0.0438</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 5</td>
<td>1.0600</td>
<td>0.9848</td>
<td>3.1901</td>
<td>245.76</td>
<td>0.0451</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Panel B: Maximum Log-likelihood values, and summary of state variables**

<table>
<thead>
<tr>
<th></th>
<th>GARCH(1,1)</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log-likelihood</td>
<td>12847.5</td>
<td>12881.1</td>
<td>12881.8</td>
<td>12883.8</td>
<td>12885.1</td>
<td>12888.4</td>
</tr>
<tr>
<td>LR p-value</td>
<td>0.0000</td>
<td>0.0056</td>
<td>0.0107</td>
<td>0.0268</td>
<td>0.0388</td>
<td>N/A</td>
</tr>
<tr>
<td>Mean of $h_z$</td>
<td>1.3473E-04</td>
<td>1.3384E-04</td>
<td>1.3037E-04</td>
<td>1.3347E-04</td>
<td>1.3186E-04</td>
<td>1.2901E-04</td>
</tr>
<tr>
<td>Std. dev. of $h_z$</td>
<td>1.2012E-04</td>
<td>1.2458E-04</td>
<td>1.1886E-04</td>
<td>1.2274E-04</td>
<td>1.2105E-04</td>
<td>1.1796E-04</td>
</tr>
<tr>
<td>Mean of $h_y$</td>
<td>N/A</td>
<td>3.9317E-06</td>
<td>6.7603E-06</td>
<td>7.8287E-06</td>
<td>9.5459E-06</td>
<td>1.1862E-05</td>
</tr>
<tr>
<td>Std. dev. of $h_y$</td>
<td>N/A</td>
<td>6.1633E-06</td>
<td>5.3816E-06</td>
<td>8.0931E-06</td>
<td>8.7196E-06</td>
<td>8.7196E-06</td>
</tr>
<tr>
<td>Mean rate of $h_y$</td>
<td>N/A</td>
<td>5.13%</td>
<td>4.93%</td>
<td>6.65%</td>
<td>7.20%</td>
<td>9.74%</td>
</tr>
</tbody>
</table>

Notes: Models 1-5 are estimated using the PF described in the paper, whereas GARCH(1,1) model is estimated using MLE. For the PF, 8192 particles are used. Panel A shows the parameter estimates of the models 1-5 and GARCH(1,1). Price of risk parameters ($\mu_z$ and $\mu_y$) are ignored by setting $\lambda_z$ and $\lambda_y$ to zero.
Table 2.4: Estimation results of equity return and volatility dynamics

Panel A: Summary of Parameter Estimates

<table>
<thead>
<tr>
<th></th>
<th>$\beta_{zj}$</th>
<th>$\beta_{yj}$</th>
<th>$\sigma_j \times 100$</th>
<th>$\rho_j$</th>
<th>$\alpha_j \times 10^6$</th>
<th>$\epsilon_j$</th>
<th>$k_j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5%</td>
<td>0.5424</td>
<td>0.4861</td>
<td>1.1500</td>
<td>0.2891</td>
<td>2.42</td>
<td>-24.34</td>
<td>-0.0022</td>
</tr>
<tr>
<td>50%</td>
<td>1.0138</td>
<td>1.0218</td>
<td>1.8502</td>
<td>0.9758</td>
<td>11.74</td>
<td>11.24</td>
<td>0.0357</td>
</tr>
<tr>
<td>95%</td>
<td>1.5268</td>
<td>1.9317</td>
<td>3.5791</td>
<td>0.9974</td>
<td>123.92</td>
<td>80.66</td>
<td>0.9621</td>
</tr>
</tbody>
</table>

Number of estimations: 509

Panel B: Correlations of Parameter Estimates

<table>
<thead>
<tr>
<th></th>
<th>$\beta_{zj}$</th>
<th>$\beta_{yj}$</th>
<th>$\sigma_j \times 100$</th>
<th>$\rho_j$</th>
<th>$\alpha_j \times 10^6$</th>
<th>$\epsilon_j$</th>
<th>$k_j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_{zj}$</td>
<td>1.000</td>
<td>0.570</td>
<td>0.455</td>
<td>0.076</td>
<td>0.095</td>
<td>0.137</td>
<td>0.112</td>
</tr>
<tr>
<td>$\beta_{yj}$</td>
<td>1.000</td>
<td>0.279</td>
<td>0.154</td>
<td>-0.044</td>
<td>0.085</td>
<td>-0.022</td>
<td></td>
</tr>
<tr>
<td>$\sigma_j \times 100$</td>
<td>1.000</td>
<td></td>
<td></td>
<td>0.096</td>
<td>0.331</td>
<td>0.117</td>
<td>0.134</td>
</tr>
<tr>
<td>$\rho_j$</td>
<td>1.000</td>
<td></td>
<td></td>
<td>-0.514</td>
<td>0.062</td>
<td>-0.747</td>
<td></td>
</tr>
<tr>
<td>$\alpha_j \times 10^6$</td>
<td></td>
<td></td>
<td></td>
<td>1.000</td>
<td>-0.052</td>
<td>0.583</td>
<td></td>
</tr>
<tr>
<td>$\epsilon_j$</td>
<td></td>
<td></td>
<td></td>
<td>1.000</td>
<td>-0.030</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_j$</td>
<td></td>
<td></td>
<td></td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel C: Significance of Parameter Correlations

<table>
<thead>
<tr>
<th></th>
<th>$\beta_{zj}$</th>
<th>$\beta_{yj}$</th>
<th>$\sigma_j \times 100$</th>
<th>$\rho_j$</th>
<th>$\alpha_j \times 10^6$</th>
<th>$\epsilon_j$</th>
<th>$k_j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_{zj}$</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.086</td>
<td>0.032</td>
<td>0.002</td>
<td>0.011</td>
</tr>
<tr>
<td>$\beta_{yj}$</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.326</td>
<td>0.055</td>
<td>0.624</td>
</tr>
<tr>
<td>$\sigma_j \times 100$</td>
<td>1.000</td>
<td></td>
<td></td>
<td>0.030</td>
<td>0.000</td>
<td>0.008</td>
<td>0.002</td>
</tr>
<tr>
<td>$\rho_j$</td>
<td>1.000</td>
<td></td>
<td></td>
<td>0.000</td>
<td>0.159</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>$\alpha_j \times 10^6$</td>
<td></td>
<td></td>
<td></td>
<td>1.000</td>
<td>0.240</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>$\epsilon_j$</td>
<td></td>
<td></td>
<td></td>
<td>1.000</td>
<td>0.504</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_j$</td>
<td></td>
<td></td>
<td></td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel D: LR test

<table>
<thead>
<tr>
<th>Significance level</th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reject restricted model</td>
<td>192</td>
<td>249</td>
<td>270</td>
</tr>
<tr>
<td>Separate beta percentage</td>
<td>37.7%</td>
<td>48.9%</td>
<td>53.0%</td>
</tr>
</tbody>
</table>

Notes: Panel A shows parameter estimates as a summary. For each parameter set of 509 estimates, values at 5%, 50% and 95% are displayed. Panel B and Panel C show the Pearson correlations and their p-values, respectively. Panel D presents loglikelihood test results with the null hypothesis that jump and diffusive betas are same, and shows number and percent of rejections over 509 estimations.
Table 2.5: Estimation results of equity options

**Panel A : Estimates of Pricing Kernel Parameters**

<table>
<thead>
<tr>
<th></th>
<th>$A_z$</th>
<th>$A_y$</th>
<th>$A_{hz}$</th>
<th>$A_{hy}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Common beta</td>
<td>Mean</td>
<td>-3.90</td>
<td>-31.28</td>
<td>629.73</td>
</tr>
<tr>
<td></td>
<td>Std. Dev.</td>
<td>6.50</td>
<td>18.26</td>
<td>1878.19</td>
</tr>
<tr>
<td>Separate betas</td>
<td>Mean</td>
<td>-3.54</td>
<td>-31.34</td>
<td>721.19</td>
</tr>
<tr>
<td></td>
<td>Std. Dev.</td>
<td>6.15</td>
<td>18.76</td>
<td>1986.62</td>
</tr>
</tbody>
</table>

**Panel B : Option Fit**

<table>
<thead>
<tr>
<th></th>
<th>RMSE</th>
<th>Log-likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td>Common beta</td>
<td>Mean</td>
<td>8.22%</td>
</tr>
<tr>
<td></td>
<td>Std. Dev.</td>
<td>3.49%</td>
</tr>
<tr>
<td>Separate betas</td>
<td>Mean</td>
<td>8.16%</td>
</tr>
<tr>
<td></td>
<td>Std. Dev.</td>
<td>3.28%</td>
</tr>
</tbody>
</table>

**Panel C : LR test**

<table>
<thead>
<tr>
<th>Significance level</th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reject restricted model (% of all firms)</td>
<td>40.9%</td>
<td>47.0%</td>
<td>49.6%</td>
</tr>
</tbody>
</table>

Notes: Panel A, B and C summarizes the results of 345 individual equity option estimations with a total number of 1,605,189 option contracts from January 3, 2000 to August 26, 2015. Given the market and individual equity parameter values estimated under physical measure, pricing kernel parameters are estimated for each set of equity options. Panel A shows the mean and standard deviations of pricing kernel parameter estimates for both model with separate betas and common beta. Panel B summarizes RMSE and log-likelihood values from 345 option estimations, whereas Panel C shows the percentages of estimations where the restricted model with common beta is rejected at 1%, 5%, and 10% significance level.
Appendix A: Model

A.1 Return innovations

Market and stock return innovations are normal shocks \((z_{t+1}, w_{i,t+1})\) and jumps \((\bar{y}_{t+1})\),

\[
\bar{r}_{m,t+1} = z_{t+1} + \bar{y}_{t+1}
\]

\[
\bar{r}_{i,t+1} = \beta_{z,i} z_{t+1} + \beta_{y,i} \bar{y}_{t+1} + w_{i,t+1}
\]

where

\[
\bar{r}_{m,t+1} \equiv r_{m,t+1} - E_t[r_{m,t+1}]
\]

\[
\bar{r}_{i,t+1} \equiv r_{i,t+1} - E_t[r_{i,t+1}].
\]

A.2 State variable dynamics

Variance contributions of normal shocks has the following Heston-Nandi type GARCH(1,1) dynamics:

\[
h_{z,t+1} = \sigma_z^2 + \rho_z (h_{z,t} - \sigma_z^2) + a_z \left( \frac{z_{t+1}^2}{h_{z,t}} - 1 - 2c_z z_{t+1} \right)
\] (2.32)

I define jump-related variance (proportional to the jump intensity) as five different but nested models, these are: (1) constant jump-variance, (2) proportional to normal-variance, (3) affine in normal-variance, (4) auto-regressive, (5) random

[Model 1] \(h_{y,t+1} = \sigma_y^2\) (2.33)

[Model 2] \(h_{y,t+1} = k h_{z,t+1}\) (2.34)

[Model 3] \(h_{y,t+1} = \sigma_y^2 + k(h_{z,t+1} - \sigma_y^2)\) (2.35)

[Model 4] \(h_{y,t+1} = \sigma_y^2 + \rho_y (h_{y,t} - \sigma_y^2) + k(h_{z,t+1} - \sigma_y^2)\) (2.36)

[Model 5] \(h_{y,t+1} = \sigma_y^2 + \rho_y (h_{y,t} - \sigma_y^2) + k(h_{z,t+1} - \sigma_y^2) + a_y \left[ \frac{y_{t+1}^2}{h_{y,t}} - (1 + \nu_y^2) - 2c_y(y_{t+1} - \nu_y y_{t+1}) \right]\) (2.37)

The variance dynamic of the idiosyncratic shocks to the stock \(i\)'s return is defined in different ways with nested models, similar to the jump-variance:

[Model 1] \(h_{i,t+1} = \sigma_i^2\) (2.39)

[Model 2] \(h_{i,t+1} = k_i h_{z,t+1}\) (2.40)

[Model 3] \(h_{i,t+1} = \sigma_i^2 + k_i(h_{z,t+1} - \sigma_i^2)\) (2.41)

[Model 4] \(h_{i,t+1} = \sigma_i^2 + \rho_i (h_{i,t} - \sigma_i^2) + k_i(h_{z,t+1} - \sigma_i^2)\) (2.42)

[Model 5] \(h_{i,t+1} = \sigma_i^2 + \rho_i (h_{i,t} - \sigma_i^2) + k_i(h_{z,t+1} - \sigma_i^2) + a_i \left( \frac{w_{i,t+1}^2}{h_{i,t}} - 1 - 2c_i w_{i,t+1} \right)\) (2.43)
A.3 Risk sensitivities: Normal and jump betas

Betas are defined as the co-variation of stock return innovations with market return innovations. Normal and jump beta can be separated,

$$
\beta_{z,i} = \frac{\text{Cov}_t(\bar{r}_{m,t+1}, \bar{r}_{i,t+1} | \bar{y}_{t+1})}{\text{Var}_t(\bar{r}_{m,t+1} | \bar{y}_{t+1})}, \quad \beta_{y,i} = \frac{\text{Cov}_t(\bar{r}_{m,t+1}, \bar{r}_{i,t+1} | z_{t+1})}{\text{Var}_t(\bar{r}_{m,t+1} | z_{t+1})},
$$

which yields that CAPM-type single beta is time-varying,

$$
\beta_{i,t} = \frac{\text{Cov}_t(\bar{r}_{m,t+1}, \bar{r}_{i,t+1})}{\text{Var}_t(\bar{r}_{m,t+1})} = \beta_{z,i} + (\beta_{y,i} - \beta_{z,i})h_{y,t}/(h_{z,t} + h_{y,t}).
$$

(2.45)

Thus, for stocks with $\beta_{y,i} > \beta_{z,i}$ ($\beta_{y,i} < \beta_{z,i}$), model implies higher (lower) exposure to market shocks when the jump-based variance increases.
Appendix B: Model details

B.1 Dynamics

Market returns:
\[ r_{m,t+1} = r_{f,t} + \left( \mu_z - \psi_N(1) \right) h_{z,t} + \left( \mu_y - \tilde{\psi}_y(1) \right) h_{y,t} + z_{t+1} + y_{t+1} \] (2.46)

Stock returns:
\[ r_{i,t+1} = r_{f,t} + \left( \mu_{z,i} - \psi_N(\beta_{z,i}) \right) h_{z,t} + \left( \mu_{y,i} - \tilde{\psi}_{y,i}(1) \right) h_{y,t} - \psi_N(1) h_{i,t} + \beta_{z,i} z_{t+1} + \beta_{y,i} y_{t+1} + w_{i,t+1} \] (2.47)

where \( \psi_N(c) \equiv \frac{c^2}{2} \) is the cumulant exponent of a standard normal shock, and
\[ z_{t+1} \sim \mathcal{N}(0, h_{z,t}), \quad w_{i,t+1} \sim \mathcal{N}(0, h_{i,t}), \quad Cov(z_{t+1}, w_{i,t+1}) = 0 \]
\[ \bar{y}_{t+1} \equiv y_{t+1} - E_t[y_{t+1}], \quad y_{t+1} \equiv \sum_{n=0}^{N_{t+1}} x_{n,t+1}, \quad x_{n,t+1} \sim \mathcal{N}(\mu_x, \sigma_x^2) \]
\[ N_{t+1} \sim \mathcal{P}(\eta_t), \quad \psi_y(c) \equiv e^{\mu_x + c^2 \sigma_x^2 / 2} - 1, \quad \tilde{\psi}_y(c) \equiv \frac{\psi_y(c)}{\mu_x^2 + \sigma_x^2} \]
\[ E_t[y_{t+1}] = \mu_x \eta_t = \nu_x h_{y,t}, \quad Var_t[y_{t+1}] = h_{y,t}, \quad E_t[\tilde{y}_{t+1}^2] = h_{y,t} + \nu_x^2 h_{y,t}^2 \]
\[ \nu_x \equiv \frac{\mu_x}{\mu_x^2 + \sigma_x^2}, \quad \nu_x^2 \equiv \frac{\sigma_x^2}{\mu_x^2 + \sigma_x^2}, \quad \eta_t \equiv \frac{h_{y,t}}{\mu_x^2 + \sigma_x^2} \]

B.2 Estimation using returns

Model dynamics can be re-written for the estimations as below.

Market returns:
\[ r_{m,t+1} = r_{f,t} + \lambda_z h_{z,t} + \lambda_y h_{y,t} + z_{t+1} + y_{t+1} \] (2.48)

Stock returns:
\[ r_{i,t+1} = r_{f,t} + \lambda_{z,i} h_{z,t} + \lambda_{y,i} h_{y,t} + \beta_{z,i} z_{t+1} + \beta_{y,i} y_{t+1} + w_{i,t+1} \] (2.49)

where
\[ \lambda_z \equiv \mu_z - \psi_N(1) = \mu_z - 1/2 \]
\[ \lambda_y \equiv \mu_y - \tilde{\psi}_y(1) = \mu_y - \left( e^{\mu_x + \sigma_x^2 / 2} - 1 \right) \]
\[ \lambda_{z,i} \equiv \mu_{z,i} - \psi_N(\beta_{z,i}) = \mu_{z,i} - \beta_{z,i}^2 / 2 \]
\[ \lambda_{y,i} \equiv \mu_{y,i} - \tilde{\psi}_{y,i}(1) = \mu_{y,i} - \left( e^{\beta_{y,i} \mu_x + \beta_{y,i}^2 \sigma_x^2 / 2} - 1 \right) \]
Appendix C: Pricing kernel

The innovations in the log-pricing kernel is defined as,

\[ m_{t+1} - E_t[m_{t+1}] = \Lambda_z z_{t+1} + \Lambda_y y_{t+1} + \Lambda_{h_z} h_{z,t+1} - E_t[h_{z,t+1}] + \Lambda_{h_y} (h_{y,t+1} - E_t[h_{y,t+1}]) \]  \hspace{1cm} (2.50)

where the expected value can be written in terms of risk-free rate, a constant and state variables,

\[ E_t[m_{t+1}] = -r_{f,t} - \alpha_0 - \alpha_{h_z} h_{z,t} - \alpha_{h_y} h_{y,t} \]

and the expected value of the state variables are

\[ E_t[h_{z,t+1}] = \sigma_z^2 + \rho_z (h_{z,t} - \sigma_z^2) \]

\[ E_t[h_{y,t+1}] = \sigma_y^2 + \rho_y (h_{y,t} - \sigma_y^2) + k (E_t[h_{z,t+1}] - \sigma_z^2) \]

The innovations in the state variables therefore are

\[ h_{z,t+1} - E_t[h_{z,t+1}] = a_z \left( \frac{z_{t+1}^2}{h_{z,t}} - 1 - 2c_z z_{t+1} \right) \]

\[ h_{y,t+1} - E_t[h_{y,t+1}] = a_y \left[ \frac{y_{t+1}^2}{h_{y,t}} - (1 + \nu_z^2 h_{y,t}) - 2c_y (y_{t+1} - \nu_z h_{y,t}) \right] + k a_z \left( \frac{z_{t+1}^2}{h_{z,t}} - 1 - 2c_z z_{t+1} \right). \]
Appendix D: Solution

D.1 Using risk-free rate

We can solve the Euler equation for $\alpha_0$, $\alpha_{hz}$ and $\alpha_{hy}$ using risk-free rate,

\[
E_t[\exp(m_{t+1} + r_{f,t})] = 1 \\
\Rightarrow \ln E_t[\exp(m_{t+1} + r_{f,t})] = 0
\]

The solution is the following.

\[
0 = \ln E_t[\exp(m_{t+1} + r_{f,t})] \\
= -\alpha_0 - \alpha_{hz}h_{z,t} - \alpha_{hy}h_{y,t} + \ln E_t\left[\exp\left(\Lambda_z z_{t+1} + \left(\Lambda_y - 2\Lambda_{hy}a_y c_y\right)\left(y_{t+1} - \nu_z h_{y,t}\right) + \left(\Lambda_{hz} a_z + \Lambda_{hy}k a_z\right)\left(\frac{z_{t+1}^2}{h_{z,t}} - 1 - 2\sigma_z z_{t+1}\right) + \Lambda_{hy}\left(a_y y_{t+1}^2 h_{y,t} - (a_y + \sigma_y^2 h_{y,t})\right)\right]\right]
\]

Thus,

\[
0 = -\alpha_0 - \alpha_{hz}/h_{z,t} - \alpha_{hy}/h_{y,t} - a_z(\Lambda_{hz} + \Lambda_{hy}k) - (\Lambda_y - 2\Lambda_{hy}a_y c_y)\nu_z h_{y,t} - \Lambda_{hy}a_y - \Lambda_{hy}a_y\nu_z^2 h_{y,t} + \psi_{z,0}(\Lambda_2) + \psi_{z,h}(\Lambda_1, \Lambda_2)h_{z,t} + \psi_{y,0}(\Lambda_4) + \psi_{y,h}(\Lambda_3, \Lambda_4)h_{y,t}
\]

where

\[
\Lambda_1 = \Lambda_z - 2\Lambda_2\sigma_z \\
\Lambda_2 = (\Lambda_{hz} + \Lambda_{hy}k)a_z \\
\Lambda_3 = \Lambda_y - 2\Lambda_4\sigma_y \\
\Lambda_4 = \Lambda_{hy}a_y
\]

Therefore the solution is,

\[
\alpha_0 = \psi_{z,0}(\Lambda_2) - \Lambda_2 + \psi_{y,0}(\Lambda_4) - \Lambda_4 \\
\alpha_{hz} = \psi_{z,h}(\Lambda_1, \Lambda_2) \\
\alpha_{hy} = \psi_{y,h}(\Lambda_3, \Lambda_4) - \Lambda_3\nu_x - \Lambda_4\nu_x^2
\]

We use the following properties of normal distribution ($z$) and distribution of normal shocks subordinated to a Poisson arrival process ($y$),

\[
E_t\left[\exp\left(p_1 z_{t+1} + p_2 z_{t+1}^2/h_{z,t}\right)\right] = \exp\left(-\ln(1 - 2p_2)/2 + \frac{p_1^2}{2(1 - 2p_2^2)} h_{z,t}\right) \tag{2.51}
\]

\[
E_t\left[\exp\left(p_1 y_{t+1} + p_2 y_{t+1}^2/\eta_t\right)\right] = \exp\left(-\ln(1 - 2p_2\sigma_x^2)/2 + \left[\exp\left(p_1^2\sigma_x^2 + 2p_1\mu_x + 2p_2\mu_x^2\right)/\left(2(1 - 2p_2\sigma_x^2)\right)\right] - 1\right) \eta_t \tag{2.52}
\]
where eq. (2.52) becomes 2.51 when the random arrival is replaced with a single deterministic arrival. It can also be written in terms of $h_{y,t}$,

\[
E_t \left[ \exp \left( \frac{p_1 y_{t+1} + p_3 y_{t+1}^2}{h_{y,t}} \right) \right] = E_t \left[ \exp \left( \frac{p_1 y_{t+1} + p_3 y_{t+1}^2}{\eta (\mu x + \sigma_x^2)} \right) \right] \\
= E_t \left[ \exp \left( \frac{p_1 y_{t+1} + \frac{p_3}{\mu x^2 + \sigma_x^2} y_{t+1}^2}{\eta} \right) \right] \\
= \exp \left( -\ln (1 - 2p_3 \nu_x^2) / 2 + \exp \left( \frac{p_1^2 \sigma_x^2 + 2p_1 \mu x + 2p_3 \nu_x \mu_x}{2(1 - 2p_3 \nu_x^2)} \right) - 1 \right) \frac{h_{y,t}}{\mu x^2 + \sigma_x^2}.
\]

I employ the following shortcuts to express relations above,

\[
\psi_z(p_1, p_2) \equiv \ln E_t \left[ \exp \left( \frac{p_1 z_{t+1} + p_2 z_{t+1}^2}{h_{z,t}} \right) \right] \\
= -\ln (1 - 2p_2) / 2 + \frac{p_1^2}{2(1 - 2p_2)} h_{z,t} \\
= \psi_{z,0}(p_2) + \psi_{z,h}(p_1, p_2) h_{z,t} \\
\tilde{\psi}_y(p_1, p_3) \equiv \ln E_t \left[ \exp \left( \frac{p_1 y_{t+1} + p_3 y_{t+1}^2}{h_{y,t}} \right) \right] \\
= -\ln (1 - 2p_3 \nu_x^2) / 2 + \exp \left( \frac{p_1^2 \sigma_x^2 + 2p_1 \mu x + 2p_3 \nu_x \mu_x}{2(1 - 2p_3 \nu_x^2)} \right) - 1 \right) \frac{h_{y,t}}{\mu x^2 + \sigma_x^2} \\
= \tilde{\psi}_{y,0}(p_3) + \tilde{\psi}_{y,h}(p_1, p_3) h_{y,t}.
\]

### D.2 Using market returns

Euler equation for market returns is,

\[
E_t[\exp(m_{t+1} + r_{m,t+1})] = 1.
\]

The solution is similar to risk-free case,

\[
0 = \ln E_t[\exp(m_{t+1} + r_{m,t+1})] \\
= \ln E_t \left[ \exp \left( -\alpha_0 + \left( \mu x - \psi_N(1) - \alpha_hz \right) h_{z,t} + \left( \mu y - \tilde{\psi}_y(1) - \alpha_hy - \Lambda y \nu_x \right) h_{y,t} \right) \right] \\
+ \Lambda_{hz} \left( h_{z,t+1} - E_t[h_{z,t+1}] \right) + \Lambda_{hy} \left( h_{y,t+1} - E_t[h_{y,t+1}] \right) + (1 + \Lambda x) z_{t+1} + (1 + \Lambda y) y_{t+1} \\
= -\alpha_0 - \Lambda_2 + \left( \mu x - \psi_N(1) - \alpha_hz \right) h_{z,t} + \left( \mu y - \tilde{\psi}_y(1) - \alpha_hy - \Lambda y \nu_x - \Lambda_4 \nu_x^2 + 2\Lambda_4 \nu_x \epsilon \nu_x \right) h_{y,t} \\
+ \ln E_t \left[ \exp \left( (1 + \Lambda x - 2\Lambda_2 \epsilon x) z_{t+1} + \Lambda_2 \frac{z_{t+1}^2}{h_{z,t}} \right) \right] + \ln E_t \left[ \exp \left( (1 + \Lambda y - 2\Lambda_4 \epsilon y) y_{t+1} + \Lambda_4 \frac{y_{t+1}^2}{h_{y,t}} \right) \right] \\
= -\alpha_0 + [\psi_{z,0}(\Lambda_2) - \Lambda_2 + \tilde{\psi}_{y,0}(\Lambda_4) - \Lambda_4] + \left[ \mu x + \psi_{z,h}(1 + \Lambda_1, \Lambda_2) - \psi_N(1) - \alpha_hz \right] h_{z,t} \\
+ \left[ \mu y + \tilde{\psi}_{y,h}(1 + \Lambda_3, \Lambda_4) - \tilde{\psi}_y(1) - \alpha_hy - \Lambda y \nu_x - \Lambda_4 \nu_x^2 + 2\Lambda_4 \nu_x \epsilon \nu_x \right] h_{y,t}.
\]

Thus,

\[
\mu x = \psi_N(1) + \psi_{z,h}(\Lambda_1, \Lambda_2) - \psi_{z,h}(1 + \Lambda_1, \Lambda_2) \\
\mu y = \tilde{\psi}_y(1) + \tilde{\psi}_{y,h}(\Lambda_3, \Lambda_4) - \tilde{\psi}_{y,h}(1 + \Lambda_3, \Lambda_4).
\]
D.3 Using stock returns

We can also solve the Euler equation using stock returns,

\[ E_t[\exp(m_{t+1} + r_{i,t+1})] = 1. \]

The solution is,

\[
0 = \ln E_t[\exp(m_{t+1} + r_{i,t+1})] \\
= -\alpha_0 - \Lambda_2 - \Lambda_4 + \left( \mu_{z,i} - \psi_N(\beta_{z,i}) - \alpha_{h_z} \right) h_{z,t} \\
+ \left( \mu_{y,i} - \tilde{\psi}_y(\beta_{y,i}) - \alpha_{h_y} - \Lambda_y \nu_x - \Lambda_4 \nu_x^2 + 2\Lambda_4 c_y \nu_x \right) h_{y,t} \\
+ \ln E_t \left[ \exp \left( (\beta_{z,i} + \Lambda z - 2\Lambda_2 c_z) z_{t+1} + 2\Lambda_2 c_z \frac{z_{t+1}^2}{h_{z,t}} \right) \right] \\
+ \ln E_t \left[ \exp \left( (\beta_{y,i} + \Lambda y - 2\Lambda_4 c_y) y_{t+1} + 2\Lambda_4 c_y \nu_x \right) \right] \\
= -\alpha_0 + [\psi_{z,0}(\Lambda_2) - \Lambda_2 + \tilde{\psi}_y(\Lambda_4) - \Lambda_4] + \left[ \mu_{z,i} + \psi_{z,h}(\beta_{z,i} + \Lambda_1, \Lambda_2) - \psi_N(\beta_{z,i}) - \alpha_{h_z} \right] h_{z,t} \\
+ \left[ \mu_{y,i} + \tilde{\psi}_y(\beta_{y,i} + \Lambda_3, \Lambda_4) - \psi_{y,h}(\beta_{y,i}) - \alpha_{h_y} - \Lambda_y \nu_x - 2\Lambda_4 c_y \nu_x \right] h_{y,t} \\
\]

Thus,

\[
\mu_{z,i} = \psi_N(\beta_{z,i}) + \psi_{z,h}(\Lambda_1, \Lambda_2) - \psi_{z,h}(\beta_{z,i} + \Lambda_1, \Lambda_2) \\
\mu_{y,i} = \tilde{\psi}_y(\beta_{y,i}) + \psi_{y,h}(\Lambda_3, \Lambda_4) - \tilde{\psi}_{y,h}(\beta_{y,i} + \Lambda_3, \Lambda_4). 
\]
Appendix E: Risk-neutralization

\[ E_t^Q \left[ \exp \left( u(r_{i,t} - E_t[r_{i,t}]) \right) \right] = E_t^Q \left[ \exp \left( u\beta_{z,i}z_{t+1} \right) \right] \times E_t^Q \left[ \exp \left( u\beta_{y,i}y_{t+1} \right) \right] \times E_t^Q \left[ \exp \left( uw_{i,t+1} \right) \right] \\
= E_t \left[ \exp \left( m_{t+1} - \ln E_t[\exp(m_{t+1})] + u(\beta_{z,i}z_{t+1} + \beta_{y,i}y_{t+1} + w_{i,t+1}) \right) \right] \\
= E_t \left[ \exp \left( \Lambda_1 z_{t+1} + \Lambda_2 z_{t+1}^2 \mu_{z,t} - \ln E_t \left[ \exp \left( \Lambda_1 z_{t+1} + \Lambda_2 z_{t+1}^2 \mu_{z,t} \right) \right] + u\beta_{z,i}z_{t+1} \right) \right] \\
\times E_t \left[ \exp \left( \Lambda_3 y_{t+1} + \Lambda_4 y_{t+1}^2 \mu_{y,t} - \ln E_t \left[ \exp \left( \Lambda_3 y_{t+1} + \Lambda_4 y_{t+1}^2 \mu_{y,t} \right) \right] + u\beta_{y,i}y_{t+1} \right) \right] \\
\times E_t \left[ \exp \left( uw_{i,t+1} \right) \right] \\
\]

Therefore,

\[ E_t^Q \left[ \exp \left( u\beta_{z,i}z_{t+1} \right) \right] \times E_t^Q \left[ \exp \left( u\beta_{y,i}y_{t+1} \right) \right] \times E_t^Q \left[ \exp \left( uw_{i,t+1} \right) \right] \\
= E_t \left[ \exp \left( \Lambda_1 + u\beta_{z,i}z_{t+1} + \Lambda_2 z_{t+1}^2 \mu_{z,t} \right) - \psi_{z,0}(\Lambda_2) + \psi_{z,h}(\Lambda_1, \Lambda_2) \right] \\
\times E_t \left[ \exp \left( \Lambda_3 + u\beta_{y,i}y_{t+1} + \Lambda_4 y_{t+1}^2 \mu_{y,t} \right) - \psi_{y,0}(\Lambda_4) + \psi_{y,h}(\Lambda_3, \Lambda_4) \right] \\
\times E_t \left[ \exp \left( uw_{i,t+1} \right) \right] \\
\]

E.1 Market normal shocks

\[ E_t^Q \left[ \exp \left( u\beta_{z,i}z_{t+1} \right) \right] = \exp \left( u\beta_{z,i}h_{z,t} + \frac{u^2 \beta_{z,i}^2}{2} \mu_{z,t} \right) \\
\exp \left( u\beta_{z,i}h_{z,t}^Q + \frac{u^2 \beta_{z,i}^2}{2} \mu_{z,t}^Q \right) = \exp \left( -\psi_{z,0}(\Lambda_2) - \psi_{z,h}(\Lambda_1, \Lambda_2)h_{z,t} \right) \\
\times E_t \left[ \exp \left( \Lambda_1 + u\beta_{z,i}z_{t+1} + \Lambda_2 z_{t+1}^2 \mu_{z,t} \right) \right] \]

\[ = \exp \left( u\beta_{z,i} \frac{(A_z - 2A_2c_z)}{1 - 2A_2} h_{z,t} + \frac{u^2 \beta_{z,i}^2}{2(1 - 2A_2)} h_{z,t} \right) \]

Thus the variance of normal shocks under risk-neutral measure is,

\[ h_{z,t}^Q = \frac{1}{1 - 2A_2} h_{z,t}. \]

We can define

\[ z_{t+1} \sim N \left( \Lambda_1 h_{z,t}^Q, h_{z,t}^Q \right), \]
which gives

\[ z_{t+1}^Q - z_{t}^Q \sim \Lambda_1 h_{z,t}^Q \]
\[ \varepsilon_{t+1}^Q - \varepsilon_{t}^Q \sim \Lambda_1 \sqrt{h_{z,t}^Q} = \tilde{\varepsilon}_{t+1} - \Lambda_1 \sqrt{h_{z,t}^Q} \]

Now we can write the risk-neutral variance dynamics using the HN-GARCH(1,1)

\[ h_{z,t+1} = w_z + b h_{z,t} + a_z \left( \frac{z_{t+1}}{h_{z,t}} - c_z \sqrt{h_{z,t}} \right)^2 \]
\[ h_{z,t}^Q = \frac{w_z}{1 - 2\Lambda_2} + b h_{z,t}^Q + a_z \left( \frac{z_{t+1}^Q}{h_{z,t}^Q} - c_z^Q \sqrt{h_{z,t}^Q} \right)^2 \]
\[ h_{z,t+1}^Q = \sigma_{z,t}^2 + \rho_z (h_{z,t} - \sigma_{z,t}^Q) + a_z \frac{z_{t+1}^Q}{h_{z,t}^Q} - 1 - 2c_z^Q \tilde{\varepsilon}_{t+1}^Q \]

where

\[ \sigma_z^2 = \frac{w_z + a_z}{1 - \rho_z} \]
\[ \rho_z = b_z + a_z c_z^2 \]

under both physical and risk-neutral measures. The relations between risk-neutral and physical parameters, therefore, are

\[ w_z^Q = \frac{w_z}{1 - 2\Lambda_2} \]
\[ a_z^Q = \frac{a_z}{(1 - 2\Lambda_2)^2} \]
\[ c_z^Q = c_z (1 - 2\Lambda_2) - \Lambda_1 \]

E.2 Market jumps

\[ E_t^Q \left[ \exp \left( u_{\beta_{y,t+1}} \right) \right] = \exp \left( u_{\beta_{y,t+1}}^Q \mu_x + u_{\beta_{y,t+1}}^2 \sigma_x^2 / 2 \right) - 1 \]
\[ = \exp \left( - \tilde{\psi}_{y,0}(\Lambda_1) - \tilde{\psi}_{y,h}(\Lambda_3, \Lambda_4) h_{y,t} + \tilde{\psi}_{y,0}(\Lambda_3) + \tilde{\psi}_{y,h}(\Lambda_3 + u_{\beta_{y,t+1}}) h_{y,t} \right) \]
\[ = \exp \left( \exp \left( \frac{u_{\beta_{y,t+1}}(\mu_x + \Lambda_3 \sigma_x^2)}{1 - 2\Lambda_4 \nu_x^2} + \frac{u_{\beta_{y,t+1}}^2 \sigma_x^2}{2(1 - 2\Lambda_4 \nu_x^2)} \right) - 1 \right) \exp \left( \frac{\Lambda_3 \sigma_x^2 + 2\Lambda_3 \mu_x + 2\Lambda_4 \nu_x \mu_x}{2(1 - 2\Lambda_4 \nu_x^2)} \right) \eta_t \]

Thus, the jump intensity and jump-variance under the risk-neutral measure is,

\[ \eta_t^Q = \exp \left( \frac{\Lambda_3^2 \sigma_x^2 + 2\Lambda_3 \mu_x + 2\Lambda_4 \nu_x \mu_x}{2(1 - 2\Lambda_4 \nu_x^2)} \right) \eta_t \]
\[ h_{y,t}^Q = \exp \left( \frac{\Lambda_3^2 \sigma_x^2 + 2\Lambda_3 \mu_x + 2\Lambda_4 \nu_x \mu_x}{2(1 - 2\Lambda_4 \nu_x^2)} \right) \frac{\mu_x^2 + \sigma_x^2}{\mu_x^2 + \sigma_x^2} h_{y,t} = \xi h_{y,t} \]
where \( \xi \equiv \exp \left( \frac{\Lambda^2 \sigma_x^2 + 2 \Lambda \sigma_x \nu_x + 2 \Lambda^2 \nu_x \sigma_x}{2(1 - 2 \Lambda \nu_x^2)} \right) \left( \mu_x^Q + \sigma_x^Q \right) \right) \left( \mu_x^2 + \sigma_x^2 \right) \) and the risk-neutral parameters are,

\[
\mu_x^Q = \frac{\mu_x + \Lambda \sigma_x^2}{1 - 2 \Lambda \nu_x^2} \\
\sigma_x^Q = \frac{\sigma_x^2}{1 - 2 \Lambda \nu_x^2}.
\]

The risk-neutral jump-variance dynamics are,

\[
h_{y,t+1} = \sigma_y^2 + \rho_y (h_{y,t} - \sigma_y^2) + k (h_{z,t+1} - \sigma_z^2) + a_y \left[ \frac{y_{t+1}^Q}{h_{y,t}} - (1 + \nu_x^2 h_{y,t}) - 2 c_y (y_{t+1} - \nu_x h_{y,t}) \right]
\]

\[
h_{Q,y,t+1}^{Q} = \sigma_y^2 + \rho_y (h_{y,t}^{Q} - \sigma_y^2) + k (h_{z,t+1} - \sigma_z^2) + a_y \left[ \frac{y_{t+1}^{Q}}{h_{y,t}^{Q}/\xi} - (1 + \nu_x^{Q} h_{y,t}^{Q}/\xi) - 2 c_y (y_{t+1}^{Q} - \nu_x^{Q} h_{y,t}^{Q}/\xi) \right]
\]

\[
h_{O,y,t+1}^{Q} = \sigma_y^2 + \rho_y (h_{y,t}^{Q} - \sigma_y^2) + k (h_{z,t+1} - \sigma_z^2) + a_y \left[ \frac{y_{t+1}^{Q}}{h_{y,t}^{Q}/\xi} - (1 + \nu_x^{Q} h_{y,t}^{Q}/\xi) - 2 c_y (y_{t+1}^{Q} - \nu_x^{Q} h_{y,t}^{Q}/\xi) \right]
\]

\[
h_{Q,y,t+1}^{Q} = \sigma_y^2 + \rho_y (h_{y,t}^{Q} - \sigma_y^2) + k (h_{z,t+1} - \sigma_z^2) + a_y \left[ \frac{y_{t+1}^{Q}}{h_{y,t}^{Q}/\xi} - (1 + \nu_x^{Q} h_{y,t}^{Q}/\xi) - 2 c_y (y_{t+1}^{Q} - \nu_x^{Q} h_{y,t}^{Q}/\xi) \right]
\]

where

\[
\sigma_y^Q = \sigma_y^2 \\
k_Q = k_Q \\
a_y^Q = a_y^2 \\
c_y^Q = c_y^Q \\
\nu_z^Q = \nu_z^Q
\]

### E.3 Idiosyncratic shocks

Idiosyncratic shocks are not systematic, and therefore there is no risk premium associated with them. This means that the risk neutral idiosyncratic shock is same as the physical one:

\[
E_t^Q \left[ \exp \left( u w_{i,t+1} \right) \right] = E_t \left[ \exp \left( u w_{i,t+1} \right) \right]
\]

\[
w_{i,t+1}^{Q} = w_{i,t+1}
\]

### Appendix F: Moment-generating function of market returns

Conditional moment-generating function (MGF) of market returns from \( t \) to \( t + \tau \) is,

\[
M_m(t, \tau) \equiv E_t^Q \left[ \exp(w r_{m,t+t+\tau}) \right] = E_t \left[ \exp(r_{f,t+t+\tau} + m_{t+t+\tau} + ur_{m,t+t+\tau}) \right]
\]
where \( r_{f,t:t+\tau-1} = \tau r_{f,t} \) since the yield curve is assumed to be flat \( (r_{f,t} = ... = r_{f,t+\tau}) \), and

\[
m_{t:t+\tau} = \sum_{j=1}^{\tau} m_{t+j}, \quad r_{m,t:t+\tau} = \sum_{j=1}^{\tau} r_{m,t+j}.
\]

We can guess that the solution is in terms of state variables, with initial conditions \( M_{m,0}(0) = 0, M_{m,z}(0) = 0, \) and \( M_{m,y}(0) = 0 \):

\[
M_{m,t}(u,\tau) = \exp \left( M_{m,0}(\tau) + M_{m,z}(\tau) h_{z,t} + M_{m,y}(\tau) h_{y,t} \right).
\]

This gives a recursive relation, once applying the law of iterated expectations,

\[
M_{m,t}(u,\tau) = E_t \left[ \exp \left( r_{f,t} + m_{t+1} + ur_{m,t+1} + (\tau - 1)r_{f,t} + m_{t+1:t+\tau} + ur_{m,t+1:t+\tau} \right) \right]
= E_t \left[ \exp \left( r_{f,t} + m_{t+1} + ur_{m,t+1} \right) E_{t+1} \left[ \exp \left( (\tau - 1)r_{f,t+1} + m_{t+1:t+\tau} + ur_{m,t+1:t+\tau} \right) \right] \right]
= E_t \left[ \exp \left( r_{f,t} + m_{t+1} + ur_{m,t+1} \right) M_{m,t+1}(u,\tau - 1) \right]
\]

The solution is below.

\[
M_{m,t}(u,\tau) = E_t \left[ \exp \left( r_{f,t} + m_{t+1} + ur_{m,t+1} + (\tau - 1)r_{f,t} + m_{t+1:t+\tau} + ur_{m,t+1:t+\tau} \right) \right]
= E_t \left[ \exp \left( r_{f,t} + m_{t+1} + ur_{m,t+1} \right) E_{t+1} \left[ \exp \left( (\tau - 1)r_{f,t+1} + m_{t+1:t+\tau} + ur_{m,t+1:t+\tau} \right) \right] \right]
= E_t \left[ \exp \left( r_{f,t} + m_{t+1} + ur_{m,t+1} + M_{m,0}(\tau - 1) + M_{m,z}(\tau - 1) h_{z,t+1} + M_{m,y}(\tau - 1) h_{y,t+1} \right) \right].
\]

Inside of the exponent expectation above is \( X_{t+1} \),

\[
X_{t+1} = r_{f,t} + m_{t+1} + ur_{m,t+1} + M_{m,0}(\tau - 1) + M_{m,z}(\tau - 1) h_{z,t+1} + M_{m,y}(\tau - 1) h_{y,t+1}
= ur_{f,t} - \alpha_0 - N_z(\tau - 1) - N_y(\tau - 1) + M_{m,0}(\tau - 1)
+ M_{m,z}(\tau - 1) \sigma_z^2 (1 - \rho_z) + M_{m,y}(\tau - 1) \sigma_y^2 (1 - \rho_y) - M_{m,y}(\tau - 1) k \rho_z \sigma_z^2
+ \left[ u \mu_z - u \psi_N(1) - \alpha h_z + P(\tau - 1) \rho_z \right] h_{z,t}
+ \left[ u \mu_y - u \psi_y(1) - \alpha h_y + M_{m,y}(\tau - 1) \rho_y - \left( \lambda_y - 2 N_y(\tau - 1) \sigma_y \right) c_y \right] \nu_x - N_y(\tau - 1) \sigma_y^2 h_{y,t}
+ \left( u + \Lambda_z - 2 N_z(\tau - 1) c_z \right) z_{t+1} + N_z(\tau - 1) \frac{z_{t+1}^2}{h_{z,t}}
+ \left( u + \Lambda_y - 2 N_y(\tau - 1) c_y \right) y_{t+1} + N_y(\tau - 1) \frac{y_{t+1}^2}{h_{y,t}}.
\]
where \( P(\tau - 1) \equiv M_{m,z}(\tau - 1) + M_{m,y}(\tau - 1)k \), \( N_z(\tau - 1) \equiv \Lambda_2 + P(\tau - 1)a_z \), and \( N_y(\tau - 1) \equiv \Lambda_4 + M_{m,y}(\tau - 1)a_y \). Now it is easy to find the exponent expectation.

\[
\ln E_t[\exp(X_{i+1})] = ur_{f,t} - \psi_{z,0}(\Lambda_2) - \tilde{\psi}_{y,0}(\Lambda_4) - P(\tau - 1)a_z - M_{m,y}(\tau - 1)a_y
+ \psi_{z,0}(N_z(\tau - 1)) + \tilde{\psi}_{y,0}(N_y(\tau - 1)) + M_{m,0}(\tau - 1)
+ M_{m,z}(\tau - 1)\sigma_z^2(1 - \rho_z) + M_{m,y}(\tau - 1)\sigma_y^2(1 - \rho_y) - M_{m,y}(\tau - 1)k\rho_z\sigma_z^2
+ \left[u\mu_z - u\tilde{\psi}_N(1) - \psi_{z,h}(\Lambda_1, \Lambda_2) + P(\tau - 1)\rho_z
+ \psi_{z,h}(u + \Lambda_z - 2N_z(\tau - 1)c_z, N_z(\tau - 1))\right]h_{z,t}
+ \left[u\mu_y - u\tilde{\psi}_h(1) - \tilde{\psi}_{y,h}(\Lambda_3, \Lambda_4) + M_{m,y}(\tau - 1)\rho_y
+ \tilde{\psi}_{y,h}(u + \Lambda_y - 2N_y(\tau - 1)c_y, N_y(\tau - 1)) \right]
+ \left(2N_y(\tau - 1)c_y - 2\Lambda_{y,h}a_yc_y\right)\nu_x - \left[N_y(\tau - 1) - \Lambda_4\nu_x^2\right]h_{y,t}
\]

Therefore,

\[
M_{m,0}(\tau) = M_{m,0}(\tau - 1) + ur_{f,t} - \psi_{z,0}(\Lambda_2) - \tilde{\psi}_{y,0}(\Lambda_4) + \psi_{z,0}(N_z(\tau - 1)) + \tilde{\psi}_{y,0}(N_y(\tau - 1))
+ M_{m,0}(\tau - 1)\left[\sigma_z^2(1 - \rho_z) - a_z\right] + M_{m,y}(\tau - 1)\left[\sigma_y^2(1 - \rho_y) - a_y - k\rho_z\sigma_z^2 - ka_z\right]
\]

\[
M_{m,z}(\tau) = M_{m,z}(\tau - 1)\rho_z + M_{m,y}(\tau - 1)k\rho_z + u\mu_z - u\tilde{\psi}_N(1) - \psi_{z,h}(\Lambda_1, \Lambda_2)
+ \psi_{z,h}(u + \Lambda_z - 2N_z(\tau - 1)c_z, N_z(\tau - 1))
\]

\[
M_{m,y}(\tau) = M_{m,y}(\tau - 1)\left[\rho_y + a_y(2c_y - \nu_x)\nu_x\right] + u\mu_y - u\tilde{\psi}_h(1) - \tilde{\psi}_{y,h}(\Lambda_3, \Lambda_4)
+ \psi_{y,h}(u + \Lambda_y - 2N_y(\tau - 1)c_y, N_y(\tau - 1))
\]

We can also write the MGF of price (instead of return),

\[
\mathcal{M}_{m,t}(u, \tau) \equiv E_t^Q[\exp(u r_{m,t+\tau})] = E_t^Q[\exp(u p_{m,t+\tau})] \exp(-up_{m,t})
\Rightarrow E_t^Q[\exp(up_{m,t+\tau})] = \mathcal{M}_{m,t}(u, \tau) \exp(up_{m,t})
\]

where \( r_{m,t+\tau} \equiv \ln(P_{m,t+\tau}/P_m) = p_{m,t+\tau} - p_{m,t} \).

### Appendix G: Moment-generating function of stock returns

Conditional moment-generating function (MGF) of stock \( i \)'s returns from \( t \) to \( t + \tau \) is,

\[
\mathcal{M}_{i,t}(u, \tau) \equiv E_t^Q[\exp(ur_{i,t+\tau})] = E_t[\exp(r_{f,t+\tau-1} + m_{t:t+\tau} + ur_{i,t+\tau})]
\]

where \( r_{f,t+\tau-1} = \tau r_{f,t} \) since the yield curve is assumed to be flat \( r_{f,t} = \ldots = r_{f,t+\tau} \), and

\[
m_{t:t+\tau} = \sum_{j=1}^{\tau} m_{t+j}, \quad r_{i,t:t+\tau} = \sum_{j=1}^{\tau} r_{i,t+j}.
\]

We can guess that the solution is in terms of state variables, with initial conditions \( M_{i,0}(0) = 0 \), \( M_{i,z}(0) = 0 \), \( M_{i,y}(0) = 0 \), and \( M_{i,i}(0) = 0 \):

\[
\mathcal{M}_{i,t}(u, \tau) = \exp \left( M_{i,0}(\tau) + M_{i,z}(\tau)h_{z,t} + M_{i,y}(\tau)h_{y,t} + M_{i,i}(\tau)h_{i,t} \right),
\]
This gives a recursive relation, once applying the law of iterated expectations,

\[ M_{i,t}(u,\tau) = E_t \left[ \exp \left( r_{f,t} + m_{t+1} + w_{i,t+1} + (\tau - 1)r_{f,t} + m_{t+1}(t+\tau) + w_{i,t+1}(t+\tau) \right) \right] \]

\[ = E_t \left[ \exp \left( r_{f,t} + m_{t+1} + w_{i,t+1} \right) E_{t+1} \left[ \exp \left( (\tau - 1)r_{f,t+1} + m_{t+1}(t+\tau) + w_{i,t+1}(t+\tau) \right) \right] \right] \]

\[ = E_t \left[ \exp \left( r_{f,t} + m_{t+1} + w_{i,t+1} \right) M_{i,t+1}(u,\tau - 1) \right]. \]

To solve the equation below,

\[ M_{i,t}(u,\tau) = E_t \left[ \exp \left( r_{f,t} + m_{t+1} + w_{i,t+1} + M_{i,0}(\tau - 1) + M_{i,z}(\tau - 1)h_{z,t+1} + M_{i,y}(\tau - 1)h_{y,t+1} + M_{i,i}(\tau - 1)h_{i,t+1} \right) \right], \]

we can first collect the terms inside of the exponent expectation (say, \( X_{t+1} \)).

\[
X_{t+1} = r_{f,t} + m_{t+1} + w_{i,t+1} + M_{i,0}(\tau - 1) + M_{i,z}(\tau - 1)h_{z,t+1} + M_{i,y}(\tau - 1)h_{y,t+1} + M_{i,i}(\tau - 1)h_{i,t+1}
\]

\[
= ur_{f,t} - \alpha_0 - N_{z,i}(\tau - 1) + N_{y,i}(\tau - 1) + M_{i,i}(\tau - 1) \left[ \sigma_i^2(1 - \rho_i) - a_i - k_i\rho_s\sigma_z^2 \right] + M_{i,0}(\tau - 1) + M_{i,z}(\tau - 1)\sigma_2^2(1 - \rho_z) + M_{i,y}(\tau - 1) \left[ \sigma_y^2(1 - \rho_y) - k\rho_s\sigma_z^2 \right] + \left[ u_{\mu z,i} - u\psi_N(\beta_{z,i}) - \alpha_{hz} + P_{\tau}(\tau - 1)\rho_z \right] h_{z,t}
\]

\[
+ \left[ u_{\mu y,i} - u\psi_N(\beta_{y,i}) + \alpha_{hy} + M_{i,y}(\tau - 1)\rho_y - \left( A_y - 2N_{y,i}(\tau - 1)\nu_y - N_{y,i}(\tau - 1)\nu_z^2 \right) \right] h_{y,t}
\]

\[
+ \left[ u - 2M_{i,i}(\tau - 1)a_i + \left[ u\psi_{N}(1) - \psi_{N}(\beta_{z,i}) - \alpha_{hz} + P_{\tau}(\tau - 1)\rho_z \right] h_{z,t}
\]

\[
+ \left[ u_{\psi_{y,i}} + \alpha_{hy} + M_{i,y}(\tau - 1)\rho_y + \psi_{y,h} \left( u_{\psi_{y,i}} + A_y - 2N_{y,i}(\tau - 1)\nu_y - N_{y,i}(\tau - 1) \right) \right] h_{y,t}
\]

\[
+ \left[ M_{i,i}(\tau - 1)\rho_i - u\psi_N(1) + \psi_{y,h} \left( u - 2M_{i,i}(\tau - 1)a_i \right) \right] h_{i,t}
\]

where \( P_{\tau}(\tau - 1) \equiv M_{i,z}(\tau - 1) + M_{i,y}(\tau - 1)k + M_{i,i}(\tau - 1)k_i, N_{z,i}(\tau - 1) \equiv \Lambda_2 + P_{\tau}(\tau - 1)a_i, N_{y,i}(\tau - 1) \equiv \Lambda_4 + M_{i,y}(\tau - 1)a_y \). Next, we find the exponent expectation,

\[
\ln E_t[\exp(X_{t+1})] = ur_{f,t} - \alpha_0 - N_{z,i}(\tau - 1) + \psi_{z,0}(N_{z,i}(\tau - 1)) - N_{y,i}(\tau - 1) + \psi_{y,0}(N_{y,i}(\tau - 1))
\]

\[
+ \psi_{w,0} \left( M_{i,i}(\tau - 1)a_i \right) + M_{i,i}(\tau - 1) \left[ \sigma_i^2(1 - \rho_i) - a_i - k_i\rho_s\sigma_z^2 \right] + M_{i,0}(\tau - 1)
\]

\[
+ M_{i,z}(\tau - 1)\sigma_2^2(1 - \rho_z) + M_{i,y}(\tau - 1) \left[ \sigma_y^2(1 - \rho_y) - k\rho_s\sigma_z^2 \right] + \left[ u_{\mu z,i} - u\psi_N(\beta_{z,i}) - \alpha_{hz} + P_{\tau}(\tau - 1)\rho_z \right] h_{z,t}
\]

\[
+ \left[ u_{\mu y,i} - u\psi_N(\beta_{y,i}) + \alpha_{hy} + M_{i,y}(\tau - 1)\rho_y + \psi_{y,h} \left( u_{\psi_{y,i}} + A_y - 2N_{y,i}(\tau - 1)\nu_y - N_{y,i}(\tau - 1) \right) \right] h_{y,t}
\]

\[
+ \left[ u - 2M_{i,i}(\tau - 1)a_i \right] h_{i,t}
\]
where \( \psi_{w,0}(x) = \psi_{z,0}(x) \), and \( \psi_{w,h}(x) = \psi_{z,h}(x) \). Therefore,

\[
M_{i,0}(\tau) = M_{i,0}(\tau - 1) + ur_{f,i} + \psi_{w,0}
\begin{pmatrix}
M_{i,i}(\tau - 1)a_i + \psi_{z,0}
\begin{pmatrix}
N_{z,i}(\tau - 1)
\end{pmatrix}
+ \bar{\psi}_{y,0}
\begin{pmatrix}
N_{y,i}(\tau - 1)
\end{pmatrix}
- \psi_{z,0}(\Lambda_2) - \bar{\psi}_{y,0}(\Lambda_4) + M_{i,i}(\tau - 1)
\begin{pmatrix}
\sigma_i^2(1 - \rho_i) - a_i - k_i(\rho_z)\sigma_z^2 + a_z
\end{pmatrix}
+ M_{i,z}(\tau - 1)
\begin{pmatrix}
\sigma_z^2(1 - \rho_z) - a_z + M_{i,y}(\tau - 1)
\end{pmatrix}
\begin{pmatrix}
\sigma_y^2(1 - \rho_y) - a_y - k(\rho_z)\sigma_z^2 + a_z
\end{pmatrix}
\end{pmatrix}.
\]

\[
M_{i,z}(\tau) = P_i(\tau - 1)\rho_z + u\mu_z - u\psi_N(\beta_{z,i}) - \psi_{z,h}(\Lambda_1, \Lambda_2) + \psi_{z,h}(u\beta_{z,i} + \Lambda_z - 2N_{z,i}(\tau - 1)c_z, N_{z,i}(\tau - 1))
\]

\[
M_{i,y}(\tau) = M_{i,y}(\tau - 1)[\rho_y + a_y(2c_y - v_x)\nu_x] + u\mu_{y,i} - u\bar{\psi}_{y,0}(\beta_{y,i}) - \bar{\psi}_{y,h}(\Lambda_3, \Lambda_4)
\]

\[
M_{i,i}(\tau) = M_{i,i}(\tau - 1)\rho_i - u\psi_N(1) + \psi_{w,h}(u - 2M_{i,i}(\tau - 1)a_i c_i, M_{i,i}(\tau - 1)a_i).
\]

We can also write the MGF of price (instead of return),

\[
M_{i,t}(u, \tau) \equiv \mathbb{E}^Q_t[\exp(ur_{i,t+t})] = \mathbb{E}^Q_t[\exp(up_{i,t+t})] \exp(-up_{i,t})
\]

\[
\Rightarrow \mathbb{E}^Q_t[\exp(up_{i,t+t})] = M_{i,t}(u, \tau) \exp(up_{i,t})
\]

where \( r_{i,t+t} = \ln(P_{i,t+t}/P_{i,t}) = p_{i,t+t} - p_{i,t} \).
Appendix H: Option prices

Following Lewis (2000), price of a call option with strike $X$ maturing in $\tau$-periods written on index level, $P_{m,t+\tau}$, can be expressed with Fourier transform of the payoff function, and the call price can be found as,

$$\text{Call}_{j,t}(X, \tau) = \mathbb{E}_t^Q[(e^{P_{m,t+\tau}} - X)^+]$$

$$= -\frac{1}{2\pi} \int_{i\kappa_{\text{L}}}^{i\kappa_{\text{U}}} \mathbb{M}_{m,t}(-i\kappa, \tau) \exp(-i\kappa p_{m,t}) \frac{X^{i\kappa + 1}}{k^2 - i\kappa} \, dk$$

$$= -\frac{X}{\pi} \int_{0}^{\infty} \text{Re} \left[ \frac{\exp \left( M_{m,0}(\tau) + M_{m,z}(\tau) h_{z,t} + M_{m,y}(\tau) h_{y,t} + i\kappa(x - p_{m,t}) \right) \right] \frac{\kappa^2}{k^2 - i\kappa} \, dk.$$ 

where $x \equiv \ln X$, and $p_{m,t} \equiv \ln P_{m,t}$. The integral above is defined for $1 < \kappa_{\text{L}} < B$, where the lower bound is also the lower bound of the strip of regularity for payoff transform $X^{i\kappa + 1}/(k^2 - i\kappa)$, and $B$ is the upper bound of the strip of regularity for the fundamental transform (i.e. $\mathbb{M}_{j,t}(-i\kappa, \tau) \exp(-i\kappa p_{m,t})$, complex-valued MGF). Payoff transform for put options is same as that for call, but the strip of regularity is $\kappa_{\text{L}} < 0$ so that the integration contour contains both poles (whereas $k_{\text{L}} > 1$ for call options to exclude the poles at $k = 0$ and $k = i$).
Bibliography


