INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.
Abstract

Essays in Financial Economics
Sze Wah Sam Cheung

The first chapter of this dissertation examines continuous-time one-factor and two-factor stochastic volatility models incorporating jumps in returns and volatility using jointly the time-series of returns and option prices on S&P 500 from 1986 to 2006. The goal of the paper is to examine the time-series of option prices. The second paper, joint with Michael Johannes, Arthur Korteweg, and Nick Polson, provides a study of the underlying structure of common asset pricing factors that are pervasively used in models of the cross-section of equity returns. The third chapter, joint with Suresh Sundaresan, develops a model of micro loans, which incorporates a) the absence of access to physical collateral, b) peer monitoring, c) threat of punishment upon default, and d) costly monitoring by lenders.
Contents

1 An Empirical Analysis of Joint Time-Series of Returns and the Term-Structure of Option Prices 1
  1.1 Introduction ............................................... 1
  1.2 Literature Review ........................................... 5
  1.3 The Model .................................................. 11
    1.3.1 Discussion ............................................. 14
    1.3.2 Option Pricing ......................................... 15
    1.3.3 Estimation Method ..................................... 16
  1.4 Empirical Results .......................................... 21
    1.4.1 Identification ......................................... 22
    1.4.2 Data Description ....................................... 22
    1.4.3 One-Factor Model ....................................... 23
    1.4.4 Two-Factor Model ....................................... 25
    1.4.5 Periods of Market Stress: Jumps and Volatility ....... 27
    1.4.6 In-sample Pricing Errors ................................ 29
    1.4.7 Risk Premia ............................................. 29
    1.4.8 Out-of-sample Pricing Performance ..................... 31
  1.5 Conclusions ................................................ 32
  1.6 Appendix .................................................. 34
    1.6.1 ATM Implied Volatility Approximation .................. 34
    1.6.2 Simulation Study ....................................... 35
2 Shocks to Asset Pricing Factors: Understanding the Risks and Realized Returns of Common Asset Pricing Factors

2.1 Introduction .............................................. 56
2.2 Time series models for return decompositions .................. 62
  2.2.1 Parameter and state estimation .......................... 64
  2.2.2 Model comparison ................................ 65
2.3 Data and empirical results .................................. 67
  2.3.1 Summary statistics and initial observations .............. 67
  2.3.2 Parameter estimates .................................. 69
  2.3.3 Volatility, jump time, and jump size estimates ......... 73
2.4 Model comparison ......................................... 74
2.5 Accounting for market returns ................................ 75
2.6 Conclusions ............................................... 76
2.7 Appendix: Algorithmic details ................................ 77
  2.7.1 MCMC algorithms .................................... 77

3 Lending Without Access to Collateral: A Theory of Micro-Loan Borrowing Rates & Defaults

3.1 Introduction .............................................. 90
3.2 Model Specification ....................................... 98
3.3 Borrower's Problem and Endogenous Default ................. 101
  3.3.1 Equilibrium & Endogenous Borrowing Rates ............. 103
  3.3.2 Role of Lender Monitoring & Defaults ................... 106
  3.3.3 Impact of Access to Physical Collateral ................. 107
  3.3.4 Competition versus Monopoly .......................... 109
  3.3.5 Contractual Features ................................. 110
  3.3.6 Fixed Rate Loan vs. Floating Rate Loan ................. 114
List of Figures

1.1 Term-structure of Implied Volatility on representative days and the average term-structure implied volatility over the sample 1996 to 2006 46
1.2 Residuals of returns equation from fitting the two-factor ............ 47
1.3 Time-series Estimates of latent processes of the one-factor SVCJ model using MCMC ........................................... 51
1.4 Time-series Estimates of latent processes of the two-factor SVCJ model using MCMC ........................................... 52
1.5 Time-series Estimates of latent processes for the simulation study using returns only ........................................ 53
1.6 Time-series Estimates of latent processes for the simulation study using jointly returns and one short-dated option .................. 54
1.7 Time-series Estimates of latent processes for the simulation study using jointly returns, one short-dated option and one long-dated option . 55
2.1 Histograms for standardized factor returns ......................... 85
2.2 Time series of estimated volatilities: Market and HML ............ 86
2.3 Time series of estimated volatilities: SMB and UMD ............ 87
2.4 Time series of estimated likelihood ratio: Market and HML .... 88
2.5 Time series of estimated likelihood ratio: SMB and UMD .... 89
3.1 Borrower’s value as a function of monitoring cost (x) and punishment cost (K) in the case of a perpetual and finite-maturity loan .... 127
3.2 Effect of two market structures: monopolistic and competitive lending
  as punishment cost and joint-liability changes .......................... 127
3.3 Effect of two market structures: monopolistic and competitive lending
  as monitoring cost and interest rate changes ............................. 128
3.4 Effect of two market structures: monopolistic and competitive lending
  as a function of maturity ................................................. 128
3.5 Delta function as a function of monitoring cost and joint-liability ... 128
3.6 Equilibrium behavior as a function of administrative cost (x) and joint-
  liability (y) when there is access to technology upon default .......... 129
3.7 Equilibrium behavior as joint-liability increases. ...................... 129
3.8 Equilibrium behavior as administrative cost increases. ............. 129
3.9 Equilibrium behavior as lender’s cost of funding increases. ......... 130
3.10 Equilibrium behavior as the maturity $T$ increases ................... 130
3.11 Equilibrium behavior as maturity $T$ increases ...................... 130
List of Tables

1.1 Factor loadings from Principal Component Analysis 45
1.2 Variance decomposition from Principal Component Analysis 45
1.3 Summary Statistics of S&P 500 Returns 45
1.4 Summary Statistics of VIX Index 46
1.5 Summary Statistics of long-dated option implied volatilities 46
1.6 S&P 500 Parameter Estimates of One-factor Volatility Model with Two Options for the sample period 1986-2006 47
1.7 S&P 500 Parameter Estimates of Two-factor Volatility Model with Two Options for the sample period 1986-2006 48
1.8 In-sample pricing error for the sample period 1986-2006 48
1.9 Risk premia for the sample period 1986-2006 49
1.10 Out-of-sample pricing error for the sample period 1996-2006 49
1.11 Simulation Study for SVCJ model using various inputs 50

2.1 Summary statistics for asset pricing factors 81
2.2 Parameter estimates for the market and SMB 82
2.3 Parameter estimates for the HML and momentum 83
2.4 Volatility correlations, in levels and in first differences for the SVCJ model for each factor 83
2.5 Parameter estimates for SMB, HML and momentum taking accounting for market returns 84

3.1 A decomposition of the Micro-Loan Market by regions 126
3.2 A decomposition of the Micro-Loan Market by lenders ......... 126
Acknowledgements

I am very thankful to my parents for their unconditional support. Their encour­agements have been very important throughout my academic development. Their willingness to support my educational desire cannot be described by words. I would also like to thank my god parents for their love and guidance throughout my life.

I am grateful to Steven Shreve for his guidance during my undergraduate and graduate studies at Carnegie Mellon University. His founding of the Bachelor of Computational Finance program started of as a small program accepting about 10 people per class. I was fortunate enough to be accepted in the second graduate class of the program. Steve’s effort in the program can clearly be seen from the placements that graduates have after graduation. Its unique design of coursework place students in good position for both work and academic development. He is truly a mentor to me. I also thank Dmitry Kramkov for his course Introduction to Mathematical Finance, which eventually led my to become interested in Finance. My education at Carnegie Mellon University inspired to take my intellectual challenge to a next level - the PhD.

During my doctoral study at Columbia University, I was fortunate to learn from many well known professors. I am especially indebted to my advisor Michael Joh­annes. I can not find enough words to express my gratitude to him. Through the course of our three-year research collaboration he exposed me to the many facets of research. I also thank my committee members Robert Hodrick, Mikhail Chernov, Mark Broadie and Claudio Irigoyen for their invaluable help and enthusiastic guid-
ance. I also thank my classmates Yael Eithensal, Sudarshan Gururaj, Jorge Murillo, Margarita Tsoutsoura, and Frank Zhang for their company and their late night time spending with me to finish up our coursework. Despite the hard work and the time spent in pursuing my doctoral study, I am very glad that I have met life-long friends.

Finally, I would like to thank my fiancé Sara Huang. During my doctoral study, I started dating with Sara, who was then my girlfriend and lived in Vancouver, Canada. Our long-distance relationship was difficult in addition to the work load of my PhD study, but she was understanding. I truly thank her for everything.
To my parents, Ka Hung Cheung and Hau Yu Lam, and my fiancé

Sara Huang
Chapter 1

An Empirical Analysis of Joint Time-Series of Returns and the Term-Structure of Option Prices

1.1 Introduction

Existing research provides strong evidence of the importance of multiple risk factors driving aggregate stock returns. In addition to the usual Brownian shocks in prices, these factors include:

- Price jumps that capture events like the crash of 1987
- Brownian shocks that capture low-frequency time-variation in the conditional volatility of returns
- Volatility jumps that capture rapid surges in conditional volatility

Understanding the nature of these risk factors and the risk premia they command is central for all of asset pricing. The extant models in the literature, while extremely flexible, have an important limitation: they have single persistent state variable, stochastic volatility, driving option prices.
The single state variable specification has a number of implications that can be viewed from different perspectives. From the perspective of the time-series of volatility, these single state variable models specify that the long-run mean of the volatility process is constant. This assumption is questioned by the long periods during which volatility is either below its long-run mean (mid 1990s and 2004-2006) or above it (1998-2002). From the perspective of option prices, the single state variable model has strong implications regarding the long-run mean and shape of the term-structure of implied volatility. For example, the implied volatility of long-dated options should be constant and if short dated volatility is above its long-run mean, the term-structure of volatility should be downward sloping.

Figure 1.1 provides evidence against the constant long-run mean specification. It shows a number of days on which short-dated volatility is well above its long-run average and the term-structure of volatility is upward sloping. All three curves start at about the same level, around 19.5%. However, the term-structure of the implied volatility on various days have very different dynamics. For instance, the one-year option is priced at 26% on June 25, 1998 while it is priced at 21% on October 8, 1997. If the true data generating process were indeed single factor, for the same set of parameters, we should see that the long-end of the term-structure from all three dates to be identical.

To formally identify the number of factors, we conduct a Principle Component Analysis (PCA) on the implied volatilities of at-the-money options on S&P 500 obtained from OptionsMetric with 30, 60, 90, 122, 152, 182, 273 and 365 days maturities. The data is available for the period January 4, 1996 to June 5, 2006. Table 1.2 reports the variance decomposition from our analysis. Our results show that the first three factors captures 97.73%, 2.07% and 0.14% each, explaining a total variation of 99.95%. The high variation captured by the first factor explains why one-factor models can generate good fits to the data. A second factor increases the total variation explained to 99.8%. This shows that a two factor model is sufficient to capture
almost all of the variation in the term-structure of at-the-money implied volatility in this sample period. We hypothesize that more factors will be needed to capture both in-the-money and out-of-the-money options.

Another interesting result is that our first PCA factor and the levels of VIX index has a correlation of 0.97 while the second PCA factor has a correlation of -0.84 with the "slope" of the implied volatilities, which is the difference of the 365 days and the 30 days implied volatility from the at-the-money options. This suggests that while most of the variation can be captured from modeling the level factor, there is an additional factor which relates to longer dated options. This can be thought of as a long memory effect in options. To capture the data, we need a persistent factor in the model. This is why we choose to model a dynamic long-run mean as the second factor in this paper as opposed to a dynamic mean-reversion, which has only transient effects.

The high correlations can also be understood from analyzing the factor loadings from our PCA. Table 1.1 reports the factor loadings. The first principal loads heavily on all maturities, with has slightly decreasing weights as maturity increases. This is a level effect on the term-structure, and is the reason why we have a very high correlation of this factor with the VIX index. The second factor loading has a negative slope. It has a positive loading (0.6506) at the short end and a negative loading (-0.4669) at the long end. The switch in sign points to information in the term-structure of the model, and hence the high negatively correlation between the slope arises. Finally, the long memory effect can also be understood by observing that all factor loadings are large for the long-dated options. In particular, the first three factor loadings for the long-dated option are 0.3035, -0.4669 and -0.4504 respectively.

Another way to interpret the need for multi-factor model is to recall that the fixed income literature has identified that at least three factor models are necessary to capture the dynamics of the yields curve. In bond pricing setting, only the maturity changes. On the other hand, in option pricing, there are two dimensions to match:
1. price/strike ratio of option and 2. term-structure of the option price. Most existing models in option pricing has only one factor and it should not be difficult to understand why they fail to capture the observed option prices.

This paper provides several contributions to the literature. First, we consider a two-factor stochastic volatility option pricing model, which is first introduced in DPS (2000) but is never estimated. The first factor is the usual stochastic volatility factor which captures the short-term deviations and spiky movements observed in option prices, whereas the second factor is a long-term component of the volatility which captures the smoothed movements in variance. Due to the overwhelming evidence of jumps, we extend this model further to accommodate for jumps in returns and volatility. Second, we develop an econometric specification which greatly increases computational feasibility in the estimation. Although this can only be used for ATM options, the method is applicable in both classical and Bayesian estimation. The technique can also be extended for other financial instruments. Third, we develop a Markov Chain Monte Carlo (MCMC) algorithm to conduct the estimation. The MCMC algorithm treats latent factors (i.e. stochastic volatility, long-run mean of stochastic volatility, jump times, jump sizes in returns and jump sizes in volatility) as missing variables and estimates them from the time-series of returns and option prices using a likelihood approach. Finally, we discuss implications of our model for risk premia and out-of-sample pricing.

In estimating the model, we use both the S&P 500 index returns and option prices. Due to the two layers of latencies in the model, returns are very insensitive to changes in the second latent factor. In simulation study, we show that even a very long series of returns alone cannot identify the second factor. Since options are informative about latent factors, we include option prices in our estimation, which also allows us to investigative the risk premia command by our model. We also show that the addition of the time-series of short-dated ATM option identifies the first factor very well, but not the second factor. Hence, to estimate our model, we propose
to use jointly the time-series of returns, a short-dated option and a long-dated option. Jumps are easily identified in all cases.

We find strong evidence of jumps in both returns and volatility in both the one-factor and two factor models, as the literature has found. Furthermore, incorporating a long-dated option identifies a second factor in volatility that is time-varying and is a smoothed version of the first factor. We show that the one-factor model cannot fit the short-dated and long-dated option simultaneously while the two-factor model can. Our estimates imply that stochastic volatility, long-run mean of stochastic volatility and jumps in returns and volatility are all important components for option pricing. This is seen in both in-sample and out-of-sample pricing analysis. Jumps in both returns and volatility significantly improve option pricing over the pure SV models while the model, while the model with jumps in returns only is clearly misspecified. The two-factor models with time-varying long-run mean greatly improve option pricing at all maturities.

The paper is organized as follows. Section 2 contains a brief literature review. Section 3 contains the model description. Section 4 describes the estimation procedure. Section 5 contains the data description and discusses the estimation results. Finally, section 6 concludes the paper.

1.2 Literature Review

Building on the seminal work by Black and Scholes (1973), many extensions of the basic model have been proposed. These include: (i) stochastic-interest rate model, (ii) jump-diffusion models, (iii) constant-elasticity-of-variance (CEV) model, (iv) stochastic volatility models, (v) stochastic volatility and stochastic interest-rates models, and (vi) stochastic volatility jump-diffusion models with and without jumps in volatility. Many of these models fall in the affine-jump diffusion models and are reviewed in the survey paper by Duffie, Pan and Singleton (2000), which we refer to as DPS (2000) hereafter. Numerical and enhanced computational feasibility allows researchers to
examine the empirical performances of these models. A partial list are: Bakshi, Cao and Chen (1997), Eraker (2004), Eraker, Johannes and Polson (2003), Jones (2003), and Pan (2002). We refer these papers in their as BCC (1997), Eraker (2004), EJP (2003), Jones (2003) and Pan (2002) respectively. These papers tend to test the models using the S&P 100 and S&P 500 index options. On the other hand, Dubinsky and Johannes (2007) propose a stochastic volatility and jump-diffusion model to incorporate earning announcements on individual equity options. We provide a brief review the papers on S&P index options individually below.

The general problem in option pricing is the inconsistency of information between option prices and underlying asset prices. Due to the asymmetry and excess kurtosis observed in returns and option data, estimation using option data alone gives implausibly high volatility-of-volatility and very negative correlations between asset return and stochastic volatility processes compared to estimation results from time-series data alone. Due to the requirement of absolute continuity across risk-neutral and physical measure, these parameters do not change. The implied volatility on call option prices closely track stochastic volatility since in one-factor models, the loading on volatility is close to unity. Consider a sample with the crash of 1987 included, estimated volatility will include many spikes. In models where jumps in volatility are excluded, the only way to capture it is by a very high volatility-of-volatility parameter. Furthermore, another general problem is that that most models underprice short maturity in-the-money options while overpricing long-dated options. To remedy this problem, many papers focus on risk premia explanations. Jump risk premia helps to explain the short maturity options as price effects from jumps respond fast in the model. On the other hand, volatility risk premia severely overprices long maturity options if they were to fit the short-dated options well. This highlights the problem in capturing the term-structure of option prices, which is the inability of existing models to capture the short-end and the long-end of implied volatilities simultaneously. BCJ (2007) provides a detailed investigation in this direction.
BCC (1997) focus on the empirical performances of various option pricing models. They evaluate the models based on three questions: (i) which is the least misspecified, (ii) which results in the lowest pricing errors, and (iii) which achieves the best hedging performance? Four sets of models are investigated: Blacks and Scholes (BS), stochastic volatility (SV), stochastic-volatility with stochastic interest-rate (SVSI) and stochastic-volatility with jumps in asset process (SVJ). Using option data S&P 500 from June 1, 1988 to May 31, 1991, they conclude that (i) incorporating stochastic volatility and jumps are important for pricing and internal consistency between option prices and the relevant time-series data, and (ii) stochastic volatility is the first-order importance in hedging performance. It is important to note that the crash October 19, 1987 with a daily change of -23% is not included in their sample. Diffusive models cannot generate such large moves without implausibly large daily volatility. The jump component is designed to capture these surprise large moves in the market. Not including the crash in the sample will certainly be in favor for the SV model as we have not seen any large movement of that magnitude since 1987. Stochastic interest-rates, on the other hand, do not significantly improve over BS or SVJ. In-sample pricing errors are as expected, in the order of SVJ, SVSI, SV, and BS. This is expected as each model has a decreasing number of parameters to fit the data. Out-of-sample pricing performances confirm that random jumps are capable of matching short-term option prices in data whereas stochastic interest rates are capable of fitting long-term options.

Pan (2002) examines the SVJ model with stochastic interest rates using joint time-series of the S&P 500 and index option prices. Her sample contains S&P 500 returns and near-the-money short-dated option from January 1989 to December 1996 at weekly frequency. She analyzes the risk premia implied from the data and concludes that jump-risk premia is significant, while volatility premia is insignificant and hard to estimate. It is interesting to note that jumps do indeed play a significant role in capturing risk premium in a period where we do not see many market crashes. Diag-
nostic test gives overwhelming evidence for long memory in volatility, which suggests the need beyond a one-factor volatility model to capture the observed dynamics. Pan (2002) argues that the primary cause why her model is incapable of capturing the dynamics between short-dated and long-dated options because of the autocorrelation generated by the model is too restrictive. The model cannot match the data implied first-step and second-step autocorrelation simultaneously since the model generates a volatility process that decays too quickly compared to data. Pricing error analysis suggests that jump-risk premia, which is only identified after including in-the-money option prices, is necessary to explain the market observed implied volatilities across strikes. Finally, the distribution of options data has heavy right tail, causing the implied volatility curve to "tip-at-the-end" and jumps in volatility will be necessary to alleviate this problem.

Similar to Pan (2002), Jones (2003) uses joint time-series of the underlying asset and option prices to estimate the CEV model and compares it with Heston's square-root stochastic-volatility model. Using S&P 100 index returns and daily implied volatilities from the options market over the sample January 2, 1986 to June 2, 2000, he concludes that the square-root SV, CEV and extended CEV models fail to generate consistent estimates between the time-series of asset prices and the cross-sections of option prices. To diagnose among the models, he computes various moments and autocorrelations implied by the models and concludes that the SV model is incapable of generating the magnitudes of negative skewness and excessive kurtosis observed in returns and implied volatilities. The CEV model is able to generate these magnitudes for the post-1987 crash sample, but still has difficulty to match the excessive kurtosis. The extended CEV model, which allows for more flexibility in the volatility-of-volatility process, is able to generate the market observed magnitudes of all four moments. Altogether, this analysis suggests that more flexible volatility processes than the SV model are needed to capture the observed moments in the time-series and the options data. Jones (2003) also considers the pricing performance
of various models and compares it with the model in EJP (2003) and concludes that jumps are necessary to capture the observed behavior of short-dated options.

Using the time-series of S&P 500 and NASDAQ index returns alone, EJP (2003) estimates the SVCJ model, which is a model with stochastic volatility and correlated jumps in both returns and volatility processes, over the sample period January 2, 1980, to December 31, 1999. This paper is the first to provide a thorough empirical study incorporating jumps in both returns and in volatility. They find that adding jumps in returns to the square-root stochastic volatility model dramatically changes the behavior of stochastic volatility. While estimated jumps in returns are infrequent, amounting to 1-2 jumps per year, they are typically large and explain 8 to 15% of the total variance of returns. On the other hand, jumps in volatility are important as they allow volatility to increase rapidly. They find that during the Fall of 1987, volatility jumped up from roughly 20% to over 50%. Once at this high level, volatility mean reverts back to its long-run level, showing the persistence of volatility. They conclude that after accounting for jumps in both returns and volatility, they find little, if any, misspecification in the model for the returns dynamics observed in data.

Eraker (2004) estimates the same model as EJP (2003) using jointly the time-series of the S&P 500 and index option prices. Due to computational burden, he limits his sample to a much shorter sample of January 1, 1987 to December 31, 1990. Consistent with Pan (2002)'s prediction, Eraker (2004) finds evidence for jumps in volatility processes. The primary reasons are the estimated clustered jumps estimated around market crashes. However, one can argue that this is because the true arrival intensity is not constant, rather than omitted jumps in volatility. Examining the estimated stochastic volatilities further shows that daily volatility spikes up on the day of crash and quickly mean reverts to "normal" levels. This speed of adjustment is not allowed in the standard square-root models. There are only two possibilities. We need either more flexible volatility processes, for example the log-volatility model of Jacquier, Polson and Rossi (1994) or jumps in volatility. However, the log-volatility, although
flexible enough to generate large instant daily changes, does not have closed-form option pricing formula, making estimation computationally infeasible using option prices. Certain types of jumps in volatilities, on the other hand, belongs to the affine class of model and has closed-form formula, allowing researchers to investigate the implied volatilities generated using their estimated parameters. Furthermore, they conclude that adding jumps in volatility is capable of generating shapes of implied volatilities that are consistent with the ones observed in the market. BCJ (2007) confirms this by analyzing the same model with various risk premia on S&P 500 futures option data. However, using jointly the options data and the time-series of S&P 500 returns, Eraker (2004) finds that adding jump components do not improve significantly over the simple Heston model. Finally, he finds evidence that the model still has problem in addressing the pricing of the options across maturities. A contrast from BCC (1997) and Pan (2002) is that Eraker (2004) has a constant interest rate.

In summary, although much work have been done in option pricing, empirical comparisons suggest that more work need to be done in reconciling the information between time-series and option prices. Existing empirical examinations concluded that (i) jumps in underlying return process capture market crashes, which were rare but catastrophic, (ii) jumps in stochastic volatility are necessary to reconcile the assumptions of the stochastic processes to data, especially with the use of option data, (iii) jump premia is necessary to capture the accuracy of option prices across moneyness, (iv) stochastic interest rate alleviates the pricing errors across option maturity, but still inadequate, (v) additional factors in the volatility process, one strongly persistent and the other quickly mean-reverting and highly volatile, are needed to provide a better fit of the third and forth conditional moments of stochastic volatility. Furthermore, mixed results are found using different sample periods of the data.
1.3 The Model

This section describes the option pricing model examined in the paper. This is a parsimonious setup as various existing models are special cases of this model. Existing paper calls for (i) jumps in underlying return process (ii) jumps in stochastic volatility, (iii) stochastic interest rate, and (iv) two-factor volatility process. Stochastic interest rates mainly alleviates the maturity dimension of option prices and for simplicity, we will assume fixed interest rate in this paper as our focus is on the additional flexibility gained from the second volatility factor. In this paper, we consider an extension of the two-factor stochastic volatility model suggested by DPS (2000), which has never been empirically investigated.

We now state our model formally. The stock price process is given under the physical $P$-measure as follows:

$$\frac{dS_t}{S_t} = \mu dt + \sqrt{V_t} dW^1_t + dJ_t$$  \hfill (1.1)

$$dV_t = \kappa_v (\theta_t - V_t) dt + \sigma_v \sqrt{V_t} dW^2_t + dJ^v_t$$ \hfill (1.2)

$$d\theta_t = \kappa_\theta (\bar{\theta} - \theta_t) dt + \sigma_\theta \sqrt{\theta_t} dW^3_t$$ \hfill (1.3)

where $W^1_t$ and $W^2_t$ are adapted and correlated Brownian motion (i.e. $\text{corr}(W^1_t, W^2_t) = \rho t$), $W^3_t$ is an adapted Brownian motion independent from $W^1_t$ and $W^2_t$, and $J_t = \sum_{i=1}^{N_t} Z_{\tau_i}$ and $J^v_t = \sum_{i=1}^{N_t} Z^v_{\tau_i}$ are jumps in the price and volatility processes respectively. Jumps are governed by $N_t$, which is a Poisson counting process for randomly arrival of jumps with constant intensity $\lambda$. The instantaneous probability of arrival of jumps is $\lambda dt$. $Z_{\tau_i}$’s are the corresponding jump sizes, which is normally distributed with mean $\mu_Z$ and standard deviation $\sigma_Z$. $Z^v_{\tau_i}$ are exponentially distributed jumps with mean $\mu_V$ in the volatility process.

Our model can be qualitatively summarized as a two-factor stochastic volatility model with contemporaneous jumps in returns and volatility. The first equation is the underlying prices process specified in returns whereas the second equation is the
volatility process with long-run mean of volatility described by the third equation.

This specification nests many of the previously analyzed models in the literature. We label the models considered in this paper:

• SV-1: The one-factor stochastic volatility model (SV) was first solved by Heston (1993) using Fourier transform method. We obtain by restricting jumps to zero (i.e. $J_t = J^0_t = 0$ for all $t$) and setting the long-run mean of volatility process to constant over time (i.e. $\sigma_0 = 0$, $\kappa_0 = 0$ and $\theta_t = \theta$ for all $t$).

• SVJ-1: The one-factor stochastic volatility model with jumps (SVJ) model is an extension to the SV-1 model allowing for random jumps to occur in stock prices. This model is obtained by restricting jumps in volatility to zero (i.e. $Z^\nu_{\tau_i} = 0$ for all jump times $\tau_i$) and setting the long-run mean of volatility process to constant over time (i.e. $\sigma_0 = 0$, $\kappa_0 = 0$ and $\theta_t = \theta$ for all $t$).

• SVCJ-1: The one-factor stochastic volatility model with contemporaneous jumps (SVCJ) in prices and volatility is an extension to the SVJ-1 model allowing for jumps in volatility. This model is obtained by setting the long-run mean of volatility process to constant over time.

• SV-2: The two-factor stochastic volatility model was introduced by DPS (2000). This model is obtained by restricting jumps to zero.

• SVJ-2: The two-factor stochastic volatility model with jumps (SVJ) model is an extension to the SV-2 model allowing for random jumps to occur in stock prices and can be obtained by restricting jumps in volatility to zero.

• SVCJ-2: The two-factor stochastic volatility model with contemporaneous jumps (SVCJ) in prices and volatility is given in Equation 1.3.

In the double jump models, we restrict jumps to arrive simultaneously in both the returns dynamics and the stochastic volatility process identification. From a state-
space perspective, this assumption reduces the model from two latent jump arrival processes to only one.

The literature in option pricing primarily focuses on one-factor stochastic volatility models. However, multi-factor extensions show great enhancements. Bates (2000) examines a two-factor volatility model where returns are governed by two volatility factors. Gallant, Hsu, and Tauchen (1999) estimate an non-affine two-factor volatility model and find evidence that the additional factor improves the goodness-of-fit for S&P 500 returns. DPS (2000) argue that the one-factor model oversimplifies the term-structure of volatility. In particular, Bollerslev and Mikkelsen (1996) argues that a one-factor volatility model decays too fast and an additional long-term memory factor is needed to capture the slow decay in volatility observed in the market. Pan (2002) also finds similar mis-specification of the one-factor volatility model due to the poor fits of the third and forth conditional moments of stochastic volatility. Her conclusions are similar to the other papers that we need two-factor volatility model where one is strongly persistent and the other factor quickly mean-revert and highly volatile.

DPS (2000) and Pan (2002) conjectures that the volatility process implied by the one-factor model decays too fast relative to the data. The autocorrelations can be computed in close-form and are

\[
corr(V_t, V_{t+\Delta}) = e^{-\kappa_v \Delta}
\]

for the one-factor model and

\[
corr(V_t, V_{t+\Delta}) = e^{-\kappa_v \Delta} + \left( e^{-\kappa_\Theta \Delta} - e^{-\kappa_v \Delta} \right) \frac{\kappa_v \sigma^2 / (\kappa_v - \kappa_\Theta)}{(\kappa_v + \kappa_\Theta) \sigma^2 / \kappa_v + \kappa_v \sigma^2 / \kappa_\Theta}
\]

for the two-factor model. Thus, the one-factor model has only one-degree of freedom via \(\kappa_v\) to match one of the many autocorrelations in the data. On the other hand, the two-factor volatility model has more degree of freedoms and thus is capable of matching more than one autocorrelation from the data. This is the mechanism in
which multi-factors volatility models capture the term-structure of the option prices.

1.3.1 Discussion

Volatility is serially correlated, as captured by the speed of mean reversion of volatility and long-run volatility \( \kappa_v \) and \( \kappa_\theta \) at long-run levels \( \theta_t \) and \( \theta_\theta \). \( \sigma_v \) is commonly referred to as volatility-of-volatility and controls the tail of the stock price distribution. In other words, \( \sigma_v \) controls the kurtosis of the model. The correlation parameter, \( \rho \), is commonly referred to as the "leverage effect" (Black (1976)) and is typically estimated to be negative. This parameter controls for the skewness of the returns distribution.

The option pricing implications of the model have been carefully examined and is briefly summarized:

1. Excess kurtosis of conditional return distribution is positive and increasing in both \( \sigma_v \) and \( |\rho| \). Positive kurtosis implies concave Black-Scholes implied volatility (IV) curves across strikes. The concavity increases in both \( \sigma_v \) and \( |\rho| \).

2. Skewness of the conditional returns distribution is positive/zero/negative for positive/zero/negative \( \rho \). As a result, negative \( \rho \) implies a downward sloping IV curve across strikes.

3. IV curves will be upward sloping in the maturity of the contract for low initial volatility \( (V_t) \) and downward sloping in the maturity of the contract for high initial volatility.

4. Jumps sizes add mass to the tails of the returns distribution. Increasing jump variances increases the left skewness of the distribution. A negative (positive) jump mean adds left (right) skewness. This implies that the implied volatility curves will be steeper than the ones in the SV model when jump mean is negative. Furthermore, the IV curve will be more concave from above for larger values of jump variances.
5. Jump in volatility component adds right skewness to the distribution of volatility, generating fatter tail for the return distribution. Additional skewness can attenuate overpricing for options that are not too far OTM, but may actually exacerbate the overpricing for far OTM calls. Hence, jumps in volatility further steepen IV curves and increase IV for ITM options.

1.3.2 Option Pricing

In this section, we will discuss how to price option using the model above. By the theory of arbitrage pricing, there exists an equivalent martingale measure, Q. Call option prices is computed as:

\[ C_t = E_t^Q [e^{-r(t-T)}(S_T - K)^+] \]

where \( E_t^Q [\cdot] \) denotes the conditional expectation up to information at time \( t \) under the measure \( Q \) and \( K \) is the strike price of the option.

Assuming that Novikov condition holds\(^1\), then the existence of a risk-neutral measure \( Q \) is guaranteed. Standard applications of Ito's lemma (see Pan (2002)) shows that the risk-neutral processes can be specified as:

\[
\begin{align*}
\frac{dS_t}{S_t} & = (r - \lambda Q \mu^Q)dt + \sqrt{V_t} dW^{Q,1}_t + \sum_{i=1}^{N^Q_t} Z^Q_{t_i} \\
\frac{dV_t}{V_t} & = (\kappa_v (\theta_v - V_t) + \eta^v)dt + \sigma_v \sqrt{V_t} dW^{Q,2}_t + \sum_{i=1}^{N^Q_t} Z^Q_{t_i} \\
\frac{d\theta_t}{\theta_t} & = \kappa_v (\theta_v - \theta_t)dt + \sigma_\theta \sqrt{\theta_t} dW^{Q,3}_t
\end{align*}
\]

where \( W^{Q,1}_t, W^{Q,2}_t \) and \( W^{Q,3}_t \) are adapted Brownian motion defined under \( Q \)-measure with the same correlation structure. \( Z^Q_{t_i} \)'s are the corresponding jump sizes defined in \( Q \)-measure, which is assumed to be normally distributed with mean \( \mu^Q_{Z} \) and standard

\(^1\) See for example, Duffie (2001)'s textbook Dynamic Asset Pricing Theory.
deviation $\sigma_z$ and $Z_{t}^{Q,v}$ are exponentially distributed jumps in the volatility process under $Q$-measure. $N_t^Q$ is a Poisson counting process with jump intensity $\lambda_t^Q$ under the risk-neutral measure. $\eta_t^v$ is the volatility risk premium, assumed to be $\eta_t^v = \eta V_t$, as most existing papers assume.

We note that risk premia only arises via a change of measure in pricing the options. Hence, volatility premia is extremely difficult to estimate even though volatility can be precisely estimated. This parameter must be identified through the term-structure of option prices. Jumps, on the other hand, do not occur frequently in a short sample, which induces identification problems. Jump premia are therefore also difficult to be identified as jumps are rare in the market. In addition, jump premia only appear in option prices and the time-series of return provides no information about them. Finally, Pan (2002) shows that jump premia can only be identified using ITM options. In contrast to the short samples of option prices used in existing paper for option prices, we hope to mitigate the identification problem by using a longer sample.

To compute option prices from the model, we must invert the characteristic function of the transition probability for stock prices. DPS (2000) provides the details to compute option prices in the affine models. Affine models have the advantage that option prices is given in semi-closed form using the numerical Fourier inversion method, providing significant computational efficiency gains over numerically solving the partial different equations. Although more efficient, this step is computational burdensome and hence, the option pricing literature usually focus on a relatively short sample. This paper provides an alternative by introducing a new method in the next section.

1.3.3 Estimation Method

This section develops a likelihood approach for estimating jump-diffusion models with stochastic volatility using Markov Chain Monte Carlo (MCMC) methods. Robert and Casella (1999) provide a general discussion of these methods, and Johannes and
Poison (2002) provide a detailed exposition of the usage of MCMC estimation in finance. EJP (2003) provides a detailed description of the application of MCMC to estimate the one-factor model. The exact posterior for the general two-factor model is given in the appendix. This approach has several advantages over classical estimation methods: (1) MCMC provides estimates of the latent volatility, long-run mean of volatility, jump times, and jump sizes; (2) MCMC accounts for estimation risk; (3) MCMC methods are shown in related settings to have superior sampling properties to competing methods; (4) MCMC methods are computationally efficient so that we can check the accuracy of the method using simulations.

To enhance computational efficiency, we use the Hull and White (1987) insight that call prices for stochastic volatility models without correlation can be approximated using the Black and Scholes formula with expected variance. They show that the approximation works reasonably well for ATM option contracts with various correlations for both short-dated and long-dated options. Chernov (2006) shows that this approximation can be extended to models with jumps, and that the approximation error is on average only 1.6%. Jones (2003) uses this method in his analysis of the joint-consistency between options and time-series data. Dubinsky and Johannes (2007) adopts this method to the their extended SV model on individual stocks accounting for earnings announcements. In simulation studies, they show that this method provides unbiased estimates. Following Chernov (2006), this is derived to be

\[ IV_{t,\tau} = \alpha + \beta_v^v V_t + \beta^\theta \theta_t \]  

(1.7)

where

\[
\beta_v^v = \frac{1}{\kappa_v^Q} (1 - e^{-\kappa_v^Q \tau}) \\
\beta^\theta = \frac{\kappa_v^Q}{(\kappa_v^Q - \kappa_\theta^Q) (\kappa_v^Q (1 - e^{-\kappa_v^Q \tau}) - \beta_v^v)} \\
\]

\[ f_t = \frac{-\gamma - \kappa_v^Q}{\kappa_v^Q} \]

We provide an alternative derivation in the Appendix.
This approximation is accurate and significantly reduces computational burden. Option pricing requires either solving a partial differential equation (PDE) or inverting the characteristic function. The implementation of either methods require numerical approximations and are computationally burdensome. Hence, using numerical approximations instead of exact prices is innocuous and is not new to the literature. However, the efficiency enhancement is a major contribution.

Approximation error can be accommodated by introducing an error term to the approximation equation:

\[
IV_{t,r} = \alpha + \beta_r^0V_t + \beta_r^\theta t + \eta_t
\]  

(1.8)

where \( \eta_t \) is an idiosyncratic error term which allows for the possibility that option prices are traded close to the model price, but not exactly equal. This can, for example, be caused by bid-ask spreads. The error distributions can be chosen to capture various sample statistics from the option data. For example, if we believe that option returns are symmetric and has no fat-tail, we would choose a normal distribution. However, since sample statistics show that option prices exhibit very negative skewness, a more skewed and fat-tailed distribution such as the t-distribution is appropriate. Before we give the formal representation, it is worth mentioning that option prices exhibit heteroscedasticity, which arises because if an asset is mispriced at time \( t \), it is also mispriced at time \( t + 1 \). In statistical terms, option pricing errors exhibit autocorrelations. Thus, we impose that the variance of option pricing errors \( \eta_t \) is proportional to \( V_t^2 \). It can be shown that \( \sqrt{\lambda_t} \epsilon_t \), where \( \lambda_t \) is distributed as inverted gamma with location and scale parameters both equal to \( \frac{\nu}{2} \) and \( \epsilon_t \sim N(0, \sigma^2 \nu V_t^2) \), is t-distributed with \( \nu \) degree of freedom. In a state space framework, this can be conveniently estimated using a Gibbs updating step. In light of this decomposition, we can view the t-distribution as an i.i.d random variable with stochastic volatility. This
added stochastic volatility component translates into excess skewness and kurtosis in the observation equation. Forcing the option pricing errors to be normally distributed will over-fit the data and worsen out-of-sample pricing.

Stochastic volatility with time-varying long-run mean models allow for different long-term dynamics of the option pricing model and long-dated derivatives are therefore expected to be very informative about the second volatility factor. Short-dated options, on the other hand, are very information about the stochastic volatility process. Thus, we expect that the loading $\beta_{t,\tau}^u \approx 1$ and $\beta_{t,\tau}^o \approx 0$ for short-dated option and $\beta_{t,\tau}^o \approx 0$ and $\beta_{t,\tau}^u \approx 1$ for long-dated option.

Before we continue, we note that in the model with a constant stochastic long-run mean of volatility, the approximation equation can be reduced to:

$$IV_{t,\tau} = \alpha + \beta_{t}^o V_t$$

(1.9)

where

$$\beta_{t}^o = \frac{1}{\kappa_0^Q} \left(1 - e^{-\kappa_0^Q \tau}\right)$$

$$\alpha = \left(\theta + \frac{\lambda^Q \mu^Q}{\kappa_0^Q}\right)(1 - \beta_v) + \lambda^Q((\mu^Q)^2 + (\sigma^Q)^2)$$

Instead of using the exact conditional distribution implied by the above process, it is more efficient to discretize the processes into daily intervals. EJP (2003) shows via simulations that the discretization bias is negligible. A first-order discretization of the model above yields:

$$IV_{t,\tau_1} = \alpha + \beta_{t,\tau_1}^o V_t + \beta_{t,\tau_1}^o \theta_t + \epsilon_t$$
$$IV_{t,\tau_2} = \alpha + \beta_{t,\tau_2}^o V_t + \beta_{t,\tau_2}^o \theta_t + \epsilon_t$$
$$y_{t+1} = \mu + \sqrt{V_t} \epsilon_{t+1} + J_{t+1} Z_{t+1}$$
$$V_{t+1} = \alpha_v + \beta_v V_t + \sigma_v \sqrt{V_t} \epsilon_{t+1} + J_{t+1}^P Z_{t+1}^P$$
where \( y_{t+1} = \ln(S_{t+1}/S_t) \) and \( J_t \) is an indicator for jumps in return and volatility. \( \epsilon_t, \epsilon^\theta_{t+1}, \) and \( \epsilon^\theta_{t+1} \) are i.i.d normal random variables in the returns equation, volatility process, and long-run mean of volatility process equations, which are the discrete time approximation for Brownian motions. Note that the first three equations are the observation equations, corresponding to short-dated ATM implied volatility, long-dated ATM implied volatility, S&P 500 returns, while the latter two are latent processes.

The above framework provides a method to estimate our model efficiently. ATM option prices is a linear function of state variables. This method can also be applied to other asset classes with a similar structure, for example, ATM bond derivative. The method is also applicable in both classical and Bayesian estimation.

Our likelihood framework also allows for the possibility of an unbalanced panel, where there are missing data for some of the dates in the sample for one series while it is available for another series. In our application, we have data for the long-dated option only for 1996-2006 while short-dated option is available for 1986-2006. Our likelihood method treats the missing data as data points which do not affect our likelihood function. We expect that latent variables are accurately estimated in the sample where full dataset is available. Structural parameters can thus be accurately learned once the latent variables are identified. Better parameter estimates will thus give more accurate estimates of latent variables in the sample with missing values. In a similar fashion, when estimating the model with returns only, we treat the first two implied volatility observation equations as missing variable and estimates the resulting state-space model.

Note that \( \alpha, \beta^\nu_{t,\tau} \) and \( \beta^\theta_{t,\tau} \) are \( Q \)-measure parameters. To estimate risk premia from this parameters, we must make further assumptions on the risk premia. There are many possible combinations of risk premia in the model. However, we can only estimate a subset of these risk premia. Equation 1.7 shows that the loading on
volatility $\beta^r$ identifies the volatility mean-reversion risk premia $\kappa_0^Q$. The loading on the long-run mean of volatility, along with $\kappa_0^Q$, identifies the long-run mean reversion risk premia $\kappa_0^Q$. As there are more parameters than equations, we can only identify one more risk premia from the remaining equation for the constant $\alpha$. As the diffusive risk premia are identified from the factor loadings, all information regarding jump risk premia are incorporated in the constant term. We can then extract one particular form of jump risk premia by assuming no risk premia on other terms. For example, to examine the jump intensity under risk-neutral measure, we hold jump mean, jump variance and volatility jump mean at their estimates under the $P$-measures. Since there are two sets of equations for the two-factor model, we extract the parameters by minimizing the sum of absolute errors. For the one-factor model, the risk premia are exactly identified once we have decided which ones to investigate. Note that this assumption means absorbing all risk premia into one parameter. This is consistent with many other papers. For example, when examining jump premia, Pan (2002) assumes only a risk premia on jump mean, and therefore her definition of jump premia includes premia on jump mean, jump variance and jump times.

### 1.4 Empirical Results

In this section, we discuss about our estimation results. We consider in-sample and out-of-sample pricing errors. A comparison will be given across maturity. Furthermore, we also examine the time-series estimates of our latent variables: stochastic volatility, long-run mean of stochastic volatility, jumps in returns and jumps in volatility. We are interested in how our model compares to other models proposed in the literature. Furthermore, we are interested in what aspects of the dynamics of time-series of returns and option prices each factor captures.
1.4.1 Identification

To examine identification of our model, we conduct a series of simulation study. A thorough discussion is in the appendix. Our results show that even a very long series of returns alone cannot identify the second factor. Due to the two layers of latencies in the model, returns are very insensitive to changes in the second latent factor. The inclusion of a short-dated option greatly enhances the identification and accurateness of estimates for stochastic volatility. An additional long-dated option is necessary to identify the second factor. Once enough data is included to impose identification, the results indicate that our procedure provides accurate inference.

1.4.2 Data Description

We estimate our model on S&P 500 as it is the most commonly analyzed equity series in the finance literature. We obtain returns data from CRSP. Due to limits on option data availability, we restrict the sample to be the same as the ones used in the options data. For options data, we use the VIX index for short-dated calls, which we obtain from Chicago Board Option Exchange (CBOE). The use of VIX avoids the numerical inversions needed for calculating IV from call options on the index. Furthermore, this gives us one data point per day for the short dated option and avoids the problem of timeliness between the closing prices of potentially illiquid options and the underlying index, as pointed out by Ait-Sahalia et al. (2001). We use the old methodology data from January 2, 1986 to June 5, 2006. This covers all the sub-samples that previous papers have analyzed. For the two-factor models, we use both the VIX and an one-year ATM option. However, due to data limitation, we only have data for the long-dated option from January 4, 1996 to June 5, 2006. The details of unbalanced estimation is in the Estimation Method section.

Table 1.3 and Table 1.4 report the sample statistics for the data set used in the paper. We also include post-crash sample statistics for comparison. Skewness and excess kurtosis drop significantly in both the returns data and in the options data.
once the crash is removed from the sample. On the return series alone in the period
November 2, 1987 - June 5, 2006, skewness dropped from -2.0596 to -0.2353 and
excess kurtosis dropped ten times once the crash is removed. The differences are
more exaggerated in the option data. Over the same period, skewness dropped from
3.1429 to 1.1270 and excess kurtosis dropped from 30.6630 to 1.5447. This difference
explains why previous researchers have found conflicting conclusions regarding the
importance of jumps versus stochastic volatility in capturing option prices.

Table 1.5 presents summary statistics of the implied volatility of various maturities
we use to conduct our out-of-sample analysis and in our PCA earlier.

1.4.3 One-Factor Model

This section discusses about the one-factor model. As our simulation studies show,
estimation using returns only cannot identify the second factor. Furthermore, since
our estimates for the one-factor model using returns only are consistent with those in
EJP (2003) and we do not report them here. In this paper, we focus on the results
estimated using jointly S&P 500 returns and options data.

Table 1.6 reports the estimates of posterior means and standard errors of the
parameters in various models within the one-factor volatility model using returns,
VIX and one-year ATM option. Parameters are quoted in annualized units, following
the convention in the option pricing literature. Since previous papers\textsuperscript{3} show that
only the model with both jumps in returns and volatility is not rejected, we will focus
our discussion in the one-factor SVCJ model in this section. We report all our results
here for completeness.

In the SVCJ model, long-run mean of volatility and speed of mean-reversion are
$\theta_v = 0.0207$ and $\kappa_v = 2.1953$. Long-run mean of volatility drops to about half of
the estimate from the SV model. This is intuitive since the SVCJ model allows
volatility to jump. The model puts less pressure on the long-run mean of volatility to

\textsuperscript{3} See BCJ (2007), EJP (2003), and Eraker (2004).
fit the high periods of volatility in the market. As with other papers, we estimate a more negative correlation after incorporating option prices. However, we have much lower estimates of volatility-of-volatility. This is primarily because of the choice of estimation technique. In the classical framework, researchers often invert volatility from option prices. Direct inversion must assume there to be no pricing error, which may be due to bid-ask spreads, on the option. Hence, the inverted volatility path will be more volatile than those that are inferred from using returns itself. Our Bayesian approach remedies this problem by inferring from the returns while assuming a pricing error on all option prices, resulting in an estimate of volatility-of-volatility which is more aligned with the time-series literature. Jump intensity is $\lambda = 0.0074$, which corresponds to about 2 jumps a year. Since volatility is precisely estimated using option data, jump times are easily identified, giving a very accurate estimate of jump intensity. Jump sizes and jump standard deviation are $\mu_Z = -1.7317\%$ and $\sigma_Z = 2.9450\%$. These estimates are similar in magnitude to the ones obtained using only the time-series of returns by EJP (2003). The time series of jumps in volatility has a mean of $\mu_V = 0.8851$ and is significant. In summary, our estimates of the SVCJ model are similar to most of the existing papers.

Throughout the paper, we mention that the one-factor model cannot capture both the short-dated and the long-dated option at the same time. We now provide formal evidence. Examining volatility of measurement error from the IV observation equations $\sigma_{T1}$ and $\sigma_{T2}$ reveal that the short-dated option ($\sigma_{T1} = 0.2978$) has a much bigger error than the long-dated option ($\sigma_{T2} = 0.0232$). The model tries to fit the long-dated option by giving up the accuracy of the short-dated option. This is not revealed in previous papers because most papers only considered options that have similar maturities. Our results indicate that we need to extend the single factor model to multi-factor specification in order to capture the term-structure of option prices. One-factor models can only fit one maturity.
1.4.4 Two-Factor Model

We now turn to the two-factor volatility model. Table 1.7 reports the posterior means and standard errors of the parameters in various models within the two-factor volatility framework estimated using returns, VIX, and the one-year ATM option. First, we note that the mean-reversion parameter of volatility in the two-factor model are higher than the one-factor model across the SV and SVCJ models, suggesting that under the two-factor model, volatility is less persistent over time. Furthermore, the mean-reversion parameters of the long-run volatility process are much lower than the mean-reversion parameters of the stochastic volatility process itself, indicating that long-run volatility is more persistent than short run volatility. To compare the long-run mean of the one-factor and two-factor stochastic processes, we note that the stationary mean of stochastic volatility in the two-factor model is \( \theta_\theta \). The long-run means of the two-factor model are uniformly higher than the one-factor model. Over the same sample period, the one-factor model keeps the long-run mean low in order to capture a long period of calm VIX movements in the market. The two-factor model estimates a very high long-run instead, capturing the upward sloping term-structure in Figure 1.1 while explaining the average movements in implied volatility. Existing literatures have demonstrated that ad-hoc adjustments to the this parameter in the one-factor model with jumps greatly improve overall price option fits, demonstrating a single factor model is incapable of simultaneously capturing the dynamics of volatility in both high and low regimes\(^4\). We also note that the volatility-of-volatility is substantially higher than their counterparts in the one-factor model than in the one-factor model while correlations are similar.

We now examine the models with jumps. The SVJ model is strongly misspecified. A positive jump mean adds skewness to the right of the conditional return distribution, while large market movements in the markets, such as the crash of 1987, tend to be negative. In particular, a positive jump mean is a consequence of the large

\(^4\) See for example, Eraker (2004)
conditional residual required on days around the 1987 crash in order to fit the VIX index. To understand why this is the case, we examine the residuals of the returns equation. Residuals from the other equations present the similar picture and are thus not reported. Figure 1.2 plots the residuals from the return equation for the two-factor SV, SVJ, and SVCJ models. Due to the large magnitudes in volatility implied by the VIX index, the residuals multiplied by volatility in the return equation is very negative. This is evident from examining the crash of 1987. Both the SV and SVJ model implies a residual of around -10%, a tailed event for a normal distribution, while the SVCJ model has a residual of -3%. Since the theoretical model implies standard normal residuals, the SVJ model requires positive jumps on those days to force the mean of the return equation to be zero. This introduces positive skewness in SVJ model, which clearly is inconsistent with the data. Thus, the SVJ model cannot simultaneously capture option prices in both low and high periods of volatility. From examining the residual plots, the SVCJ model is the only model which is not rejected.

In the SVCJ model, the estimate for jump mean and standard deviation in returns are $\mu_Z = -1.2592\%$, $\sigma_Z = 3.3503\%$, a jump intensity of $\lambda = 0.0122$ and a jump-mean of volatility of $\mu_V = 1.0432$. Jump mean, jump variance and jump mean of volatility are significant at 1% and are all lower than their counterparts in the one-factor model while jump intensity is higher. Jumps arrives about 3 times per year. Due to the movements in the long-run mean of volatility, the model estimates more frequent jumps with a smaller magnitudes.

As in the one-factor model, one way to evaluate the is by considering the volatilities of measure error. For the SVCJ model, $\sigma_{T_1} = 0.0692$ while $\sigma_{T_2} = 0.0415$. The error on the short-dated option is about only one-fifth of the counterpart in the one-factor model. Both standard deviations of the long and short dated options are small and are on similar magnitude, indicating that the two-factor model has the ability to capture both the short-dated and the long-dated options simultaneously.

To summarize, we show that only the SVCJ models within the two-factor frame-
work are consistent with the returns and options data. To understand the performances of two-factor model relative to the one-factor model, we consider pricing errors in later sections.

1.4.5 Periods of Market Stress: Jumps and Volatility

MCMC provides estimates of latent state variables, which includes stochastic volatility, long-run mean of volatility, jump times, jumps in returns and jumps in volatility. These estimates provide a mean to understand how the dynamics of these risk factors affects option price. This section analyzes how these risk factors evolve in periods of market stress for the SVCJ models.

Figure 1.3 shows the estimated spot volatilities, jump sizes in returns and jump sizes in volatility of the one-factor models. We first examine the crash of 1987. On October 19, 1987, the S&P 500 dropped 23% on a single day. On the same day, we estimate a jump in spot volatility, increasing annualized volatility from 24.57% to 62.46% with a jump in return of only -16%. Using data on returns alone, EJP (2003) estimates that jump in returns deliver a -14% return. Our smaller estimates is a consequence of the fact that that option returns are much more negatively skewed than the index returns, indicating the relative importance of jumps in volatility over jumps in returns. While many existing papers suggest that jump intensity is a function of volatility, we have not identified neighboring jumps in the crash of 1987. The large movements in returns are attributed to Brownian shocks. Volatility remains high for the entire week, and slowly decays back to the levels prior to the crash in about 6 months. We also note that this jump in volatility is the highest magnitude estimated in our sample.

In October 1997, there was a -7% move on the 27th and a 5% on the 28th in the index. VIX also increased from 20% to 38.5%. The SVCJ model attributes a jump in volatility on the 27th with a jump size in return of -5.85%. Once volatility increased, a 5% movement becomes within the 95% confidence interval and thus the following
day is driven by high volatility. Late Summer and early Fall 1998 were also periods where we observe moves in magnitude of 3-7%. On the three dates where we identified jumps, August 4, 1998, August 27, 1998 and September 30, 1998, the estimated jump sizes are only about half of the movements in their corresponding returns in the index on these dates. Therefore, we attribute that most of the variation in this period is explained by high volatility. In particular, on August 31, 1998, when the S&P 500 suffers a 7% decrease and implied volatility increased by about by 50%, we have not identified a significant jump size in return. Similar results are found for the periods September 2001 and Summer 2002 to Spring 2003. Our analysis indicates that both jumps in volatility and returns are important component of periods of market stresses and play very different roles in capturing the dynamics of the financial market.

Figure 1.4 shows the estimated spot volatilities (thin black line), long-run mean of volatility (thick red line), jump sizes in returns and jump sizes in volatility of the one-factor models. In general, the pictures are very similar to their counterparts in the one-factor model. Thus, we focus our discussion in the long-run mean of volatility. The estimates of long-run mean of volatility are visually seen to be a "smoothed" version of the stochastic volatility. For example, when volatility jumped in October 1987, the long-run mean stayed persistent to its previous levels. The two processes also track each other very closely and stochastic volatility mean-reverts to the long-run mean process at every instant in time. When stochastic volatility is above its long-run mean during the crash of 1987, it quickly drops to the long-run level subsequent to the increase. In the period 1991 to 1996, stochastic volatility is consistently below its long-run counterparts in order to compensate for the over reactive volatility process during the 1987. From 1996 to 2000, they track each other very closely. Finally, from 2000 to 2003, stochastic volatility is consistently above its long-run mean and thus is compensated by the lower estimates in the final period from 2003 to 2006. The estimates of the volatility process clearly shows that it is dynamic overtime. This is exactly the reason why the term-structures of implied volatilities like the ones in
Figure 1.1 are observed in the market and single factor models cannot capture the term-structure of options.

1.4.6 In-sample Pricing Errors

This section considers the pricing errors of our model within our sample. We measure in-sample pricing error as the square-root of mean-squared (RMSE) of differences between model-implied and market-observed Black-Scholes implied volatilities. Comparing in implied volatility unit as opposed to dollar units avoids placing undue weight on expensive options, such as deep-ITM or longer-dated options. Table 1.8 shows the resulting calculations.

The SVCJ model improves by about 46% over the SV models for the short-dated ATM option for the two-factor model, while about 11% for the one-factor model. This is intuitive as the SVCJ model has more degree of freedom to fit option prices. We also find that the in-sample pricing errors are smaller for the two-factor model than on the one-factor model, while it is in reversed direction for the long-dated option. As mentioned earlier, the standard deviation on the observation equation errors suggested that the one-factor model overfit the long-dated option, resulting in better performances of the one-factor model in pricing long-dated options than on short-dated options. The in-sample error for the long-dated option are similar in magnitudes across SV and SVCJ models.

1.4.7 Risk Premia

This section analyzes the risk premia extracted from our model. As mentioned earlier, risk premia arise only in option prices and therefore are fully incorporated into the estimates of factor loadings on stochastic volatility and long-run mean of volatility and in the constant term in the implied volatility Equation 1.7. We focus our discussion on the two-factor models.

First, we examine the diffusive risk premia. We estimate significant diffusive
volatility risk premia for all models. This parameter is identified via the term-structure of option prices. As an experiment, we tried estimating this parameter with only the short-dated option and we could not obtain significant volatility risk premia. We need an additional longer-dated option to identify this parameter. In the one-factor model, Pan (2002) suggests that the SV model relies heavily on this parameter to price long-dated options, and this parameter does not affect ITM nor OTM option pricing.

Using the standard definition of diffusive risk premia \( \eta_v = \kappa_v - \kappa_v^Q \), we find the diffusive risk premia is negative for the SV model and positive for the SVCJ model. Theoretical model does not constrain this parameter to be of either sign. Pan (2002) suggests that a negative diffusive risk premia on volatility leads to the overpricing of the SV model on long-dated options in a one-factor model. However, her argument does not hold for the two-factor model. In the two-factor model, there is another form of diffusive risk premia which arises from the long-run mean of stochastic volatility. We note that the estimate of mean-reversion of the long-run mean of stochastic volatility doubled from \( P \)-measure to \( Q \)-measure in the SV model, while it is only about 20% higher for the SVCJ model. In the absence of jumps, the SV model must increase volatility in order to capture historical option prices, while the SVCJ model adjust for this via a jump risk premia. The SV model treats high observations as frequent diffusive movements whereas the SVCJ model treats these observations as rare events.

Our estimates of jump parameters are significant under \( Q \)-measure. The significant estimates should not be a surprise as jumps cannot be perfectly hedged with a finite number of financial instruments. Our estimates are much smaller than those estimated in BCJ (2007) while in similar magnitudes to Bates (2006), who developed a filtration based maximum likelihood approach to estimate latent affine processes. We believe this is because our estimates only incorporates ATM options while jump premia affects ITM option prices the most. Therefore, an interesting exercise is to
test our estimates on ITM option prices. However, this only tests the ability our estimates to explain the cross-section of option prices, but not the model. To fully assess the model across all dimensions of option prices, we must use ITM options as inputs. However, this is not the main purpose of this paper. Our risk-neutral estimates for jump mean and jump variances are larger in magnitudes in the one-factor model, suggesting additional kurtosis is needed to fit options data. On the other hand, jump mean and jump variances are almost identical in the two-factor model across measure.

Our estimates of jump intensity and volatility jump mean present interesting differences between the one-factor and two-factor models. In the one-factor model, jump intensity is about one jump every 2-3 years under the risk-neutral measure, while the estimates in the two-factor model amounts a jump every 1-2 years. The difference arises because of long-run mean of stochastic volatility is time-varying in the two-factor model. The two-factor model places less burden on jumps since some of the movements in implied volatility is explained by movements in the long-run mean of volatility. Nevertheless, after accounting for an additional factor in stochastic volatility to capture the term-structure of option prices, jump premia remain significant. In summary, there is no substitutability between jumps and additional volatility factors. If there were, we would see either the risk premium on the long-run mean or jump risk premium to be significant, not both.

1.4.8 Out-of-sample Pricing Performance

This section analyzes the out-of-sample pricing performance of our model. We want to see how accurately our model predicts option prices. We consider pricing performances across maturities. To do this, we use parameter estimates in Table 1.6 and Table 1.7 along with risk premia given in Table 1.9 to calculate the model implied volatilities. We then compute RMSE to evaluate our model. For both models, we can compute out-of-sample pricing errors for options with maturities 60, 90, 122, 152,
182, and 273 days.

Table 1.10 displays the results. Due to the many possible combinations of risk premia discussed in the previous section, we present the one that gives the best results. We consider out-of-sample pricing errors for the case with risk premia in jump intensity. As a base case, we also include the diffusive risk premia (both mean-reversion in stochastic volatility and in long-run volatility) in the analysis.

First, we consider the one-factor model. Pricing improvement of the jump model ranges from about 2% for the 60 days option to 51% for the 273 days option. This is expected as our in-sample analysis show that the one-factor model over-weights the long-dated option while ignoring the fit of the short-dated option. On average, the model with jump premia offers about 10% pricing improvement over the pure diffusive model. In the two-factor model, the results are mixed. Pricing improvement of the jump model ranges from about 8% for the 60 days option to 30% for the 273 days option. The improvements are not as drastic as the one-factor model, but still provides a 13% improvement on average. In both models, out-of-sample pricing error decreases as maturity increases.

To compare the one-factor and two-factor model, we compare the pricing errors on a column-by-column basis. We see that the all two-factor models uniformly outperform their one-factor counterparts. The pure diffusive model has an outperformance of about 23% while the jump models outperforms by about 27%. The out-of-sample pricing analysis clearly shows the importance of an additional factor to model option prices. Our analysis is conducted on only ATM options. We expect the results to be stronger for ITM and OTM options.

1.5 Conclusions

In this paper, we use both the time-series of returns and option prices to address a number of important option pricing issues. First, we develop a two-factor stochastic volatility option pricing model with jumps in returns and volatility. Second, we
develop an econometric specification, based on the Hull-White relationship, which increases computational time in the estimation. This can be used for both classical maximum likelihood estimation or Bayesian style estimation. Third, we develop a Markov Chain Monte Carlo (MCMC) algorithm to estimate the structural parameters and the latent factors. Finally, we discuss implications of our model for risk premia and out-of-sample pricing.

Using time-series of returns, we find strong evidence of jumps in both returns and volatility. Furthermore, incorporating a long-dated option identifies a second factor, the long-run mean of stochastic volatility, that is time-varying and is a smoothed version of the first factor, the stochastic volatility. Our estimates imply that both stochastic volatility and long-run mean of stochastic volatility, along with jumps in returns and volatility, are all important components for option pricing. Jumps are important for pricing short-dated options while long-run mean of stochastic volatility is important for pricing long-dated option. This is seen in both in-sample and out-of-sample pricing analysis. Jumps in both volatility and returns are shown to have significant pricing improvement over the pure SV models in both the one-factor and two-factor models. We find significant improvements from adding jumps to returns and volatility in pricing short-dated options. The two-factor models with time-varying long-run mean of stochastic volatility provide great improvements in pricing long-dated options. Finally, we provide formal evidence that the single factor model can only fit one maturity.

From an option pricing exercise perspective, our results show that if you are only interested at accurately pricing and predicting options of only one maturity, the one-factor model with jumps is sufficient. However, if the objective is to price a panel of options with different maturities, at least two factors are necessary. If the panel contains both in-the-money and out-of-the money options in addition to ATM options, it is necessary to incorporate these options as inputs to estimate a magnitude of jump premia which is capable of pricing these options. Further research along this area is
promising.

1.6 Appendix

1.6.1 ATM Implied Volatility Approximation

This section derives the Hull-White ATM implied volatility approximation. We derive the approximation for the pure stochastic volatility case. The same methodology extends directly to the models with jump. Our goal is to find the quadratic variation of volatility, which is given by:

\[ f(t, V_t, \theta_t) = E_t \left[ \frac{1}{\tau} \int_t^T V_u du \right] \]

where the volatility process \( V_t \) is given in equation 1.3. Assume that the function \( f(t, v, \theta) \) is twice differentiable in \( t, v \) and \( \theta \). By Ito's Lemma, we have:

\[ df(t, v, \theta) = f_t dt + f_v dV_t + f_{\theta} d\theta_t + \frac{1}{2} (f_{vv}(dV_t)^2 + f_{\theta\theta}(d\theta_t)^2 + 2f_{v\theta} dV_t d\theta_t) + \frac{1}{\tau} \]

Assume that the solution to this equation is linear in state variables:

\[ f(t, v, \theta) = \alpha(t, T) + \beta^v(t, T)v + \beta^\theta(t, T)\theta \]

The relevant derivatives are:

\[ f_t = \frac{d\alpha(t, T)}{dt} + \frac{d\beta^v(t, T)}{dt}v + \frac{d\beta^\theta(t, T)}{dt}\theta \]
\[ f_v = \beta^v(t, T) \]
\[ f_{\theta} = \beta^\theta(t, T) \]
Substituting into the Ito's Lemma gives:

\[ df(t,v,\theta) = (f_t + f_v \kappa_v(\theta_t - V_t) + f_\theta \kappa_\theta(\theta - V_t))dt + f_v \sigma_v \sqrt{(V_t)}dW_t^1 + f_\theta \sigma_\theta \sqrt{(\theta_t)}dW_t^2 \]

By Feyman-Kac's formula,

\[ f_t + f_v \kappa_v(\theta - v) + f_\theta \kappa_\theta(\theta - v) = 0 \]

Substituting the derivatives into this partial differential equation and by the method of undetermined coefficient, we reduce the complicated partial differential equation into 3 separate ordinary differential equations:

\[ \beta^\theta(t,T) \kappa_\theta + \frac{d\alpha(t,T)}{dt} = 0 \]
\[ \frac{d\beta^v(t,T)}{dt} - \beta^v(t,T) \kappa_v + \frac{1}{\tau} = 0 \]
\[ \frac{d\beta^\theta(t,T)}{dt} + \beta^v(t,T) \kappa_v - \beta^\theta(t,T) \kappa_\theta = 0 \]

Using the second equation, we can solve for \( \beta^v(t,T) \) in closed-form using the method of undetermined coefficient. Substituting this into the third equation and solving the resulting ordinary differential equation yields \( \beta^\theta(t,T) \). Finally, substituting \( \beta^\theta(t,T) \) in the first equation and integrating yields \( \alpha(t,T) \). This completes the derivation.

### 1.6.2 Simulation Study

The goal of this simulation study is to examine the ability of our MCMC algorithm to estimate the general model discussed in Equation 1.3. Using the Euler discretization method discussed earlier, we sample 8000 data points from model. Table 1.11 displays the true values used in the simulation and summarizes results for the SVCJ-2 model.

When estimating the model using returns only, we impose that that estimated variance of the long-run mean process is equal to its true value. This restriction is imposed to enhance identification of the long-run mean process. However, our
algorithm has problem in estimating the SVCJ-2 model. In particular, correlation and volatility-of-volatility are very difficult to estimate, as the literature have found. Parameters governing jumps in returns are fairly accurately estimated. Due to the underestimation of the variation in stochastic volatility, excess return is significantly overestimated. Finally, the mean of jump size in return is underestimated.

The lack of identification using returns only calls for additional data in the estimation. Hence, in addition to returns, we use an additional short-dated option and display the results in the "One Option" columns. Parameters governing the first latent process are much more accurately estimated. Jump parameters for both returns and volatility are fairly accurately estimated too. The variation in the long-run mean process is, however, still not very accurate. To identify the long-run mean process, we include an additional long-date option and display the results in the "Two Options" columns. The results for the long-run mean process are now fairly accurate. We note that the coefficients for the implied volatility equation are problematic due to the persistent of volatility and long-run mean of volatility. This is analytical to the fact that risk premia is difficult to estimation since these parameters are measured under the risk-neutral measure. The error translates into an overestimation in the constant in the implied volatility equation.

To see the problems of the identification, we provide our MCMC estimates of the latent processes in Figures 1.5, 1.6, 1.7 for the SVCJ-2 model estimated using returns, one option and two options respectively. The black line represents the true process while the red line shows our MCMC estimates. In general, the first factor, volatility, is easily identified once option is incorporated. However, the second factor requires additional options to identify. Jumps are, on the other hand, easily identified in all cases.
1.6.3 MCMC Algorithm

This section derives the MCMC algorithm used in the estimation of the general two-factor model.

\[ y_{t+1} = \alpha + \beta y_t + \sqrt{V_t} \epsilon_{t+1} + J_{t+1} Z_{t+1} \]

\[ V_{t+1} = (1 - \beta_v) \theta_t + \beta_v V_t + \sigma_v \sqrt{V_t} \epsilon_{t+1} + J_{t+1} Z_{t+1}^v \]

\[ \theta_{t+1} = \alpha \theta + \beta \theta_t + \sigma_v \sqrt{\theta_t} \epsilon_{t+1}^\theta \]

Define

\[ \Sigma = \begin{pmatrix} 1 & \sigma_v \rho & 0 \\ \sigma_v \rho & \sigma_v^2 & 0 \\ 0 & 0 & \sigma_\theta^2 \end{pmatrix} = \begin{pmatrix} 1 & \psi & 0 \\ \psi & \psi^2 + \Omega & 0 \\ 0 & 0 & \sigma_\theta^2 \end{pmatrix} \]

The inverse of the variance-covariance matrix is easily calculated to be:

\[ \Sigma^{-1} = \begin{pmatrix} \frac{1}{1 - \rho^2} & -\frac{\sigma_v \rho (1 - \rho^2)}{\sigma_v^2 (1 - \rho^2)} & 0 \\ -\frac{\sigma_v \rho (1 - \rho^2)}{\sigma_v^2 (1 - \rho^2)} & \frac{1}{\sigma_v^2 (1 - \rho^2)} & 0 \\ 0 & 0 & \frac{1}{\sigma_\theta^2} \end{pmatrix} \]

The system is thus tri-variate normal, and the likelihood function is given by:

\[ p(y_{t+1}, V_{t+1}, \theta_{t+1} | V_t) = \frac{1}{\sqrt{\theta_t} V_t} |\Sigma|^{-1/2} (2\pi)^{-1} \exp \left( -\frac{1}{2} R_{t+1} \Sigma^{-1} R_{t+1}^T \right) \]

where

\[ R_{t+1} = \begin{pmatrix} \frac{y_{t+1} - \alpha - \beta y_t - \sqrt{V_t} \epsilon_{t+1} - J_{t+1} Z_{t+1}}{\sqrt{V_t}} & \frac{V_{t+1} - (1 - \beta_v) \theta_t - \beta_v V_t - J_{t+1} Z_{t+1}^v}{\sqrt{V_t}} & \frac{\theta_{t+1} - \alpha \theta - \beta \theta_t}{\sqrt{\theta_t}} \end{pmatrix} \]
Note that
\[
\sum_{t=1}^{T} R_{t+1} \Sigma^{-1} R_{t+1}^T = \text{tr}(\sum_{t=1}^{T} R_{t+1} \Sigma^{-1} R_{t+1}^T) \\
= \sum_{t=1}^{T} \text{tr}(R_{t+1} \Sigma^{-1} R_{t+1}^T) \\
= \sum_{t=1}^{T} \text{tr}(R_{t+1}^T R_{t+1} \Sigma^{-1}) \\
= \text{tr}(R^T R \Sigma^{-1})
\]

Using this, we derive the following expression, which is useful in most of the updating of the model:
\[
\text{tr}(R_{t+1}^T R_{t+1} + R_t^T R_t) \Sigma^{-1} = \\
\frac{1}{1-\rho^2}(\frac{\Sigma_{t+1}^2}{\Sigma_t} + \frac{\Sigma_t^2}{\Sigma_{t+1}}) + 2\left(-\frac{\rho}{\sigma_t(1-\rho^2)}\right)\left(\frac{\Sigma_{t+1}(V_{t+1} - (1-\beta_0)\theta_t V_t - J_{t+1} Z_{t+1}^T)}{\Sigma_t V_t} \right) + \\
\frac{r_t(V_t - (1-\beta_0)\theta_t V_t - J_t Z_t^T)^2}{\Sigma_t V_t} + \frac{1}{\sigma^2(1-\rho^2)}\left(\frac{\Sigma_{t+1}^2}{\Sigma_t} \right) + \frac{1}{\sigma^2}\left(\frac{\theta_{t+1} - \alpha \beta_{t+1}}{\theta_t} + \frac{(\theta_{t+1} - \alpha \beta_{t+1})^2}{\theta_t^2} \right)
\]

and
\[
r_{t+1} = y_{t+1} - \alpha - \beta y_t - \sqrt{\Sigma_t} \epsilon_{t+1} - J_{t+1} Z_{t+1}
\]

We now add options in the model. Using the Hull-White approximation of at-the-money options, we have:

\[
IV_{t, r_1} = \alpha IV(r_1) + \beta_{IV}^\alpha (r_1) V_t + \beta_{IV}^\theta (r_1) \theta_t + \sigma_{r_1} \epsilon_t^r_1
\]
\[
IV_{t, r_2} = \alpha IV(r_2) + \beta_{IV}^\alpha (r_2) V_t + \beta_{IV}^\theta (r_2) \theta_t + \sigma_{r_2} \epsilon_t^r_2
\]

We assume that \(\epsilon_t^r_1\) and \(\epsilon_t^r_2\) are uncorrelated and independent from the parameter in \(P\) measure. We can derive the joint likelihood of the model with options to be:

\[
p(IV_{t+1, r_1}, IV_{t+1, r_2}, y_{t+1}, V_{t+1}, \theta_{t+1}| V_t) = \frac{1}{\sqrt{\theta_t V_t}} \left| \Sigma \right|^{-1/2} (2\pi)^{\frac{1}{2}} \exp\left( -\frac{1}{2} R_{t+1} \Sigma^{-1} R_{t+1}^T \right)
\]
\[ x \exp \left( -\frac{1}{2\sigma_{\tau_1}^2} (IV_{t+1, \tau_1} - \alpha IV(\tau_1) - \beta IV(\tau_1) V_{t+1} - \beta^\theta IV(\tau_1) \theta_{t+1})^2 \right) \]
\[ \times \exp \left( -\frac{1}{2\sigma_{\tau_2}^2} (IV_{t+1, \tau_2} - \alpha IV(\tau_2) - \beta IV(\tau_2) V_{t+1} - \beta^\theta IV(\tau_2) \theta_{t+1})^2 \right) \]

**Conditional Posteriors for \( \beta_v \)**

Let

\[ y_t^v = \beta_v x_t + \sigma_v \epsilon_{t-1} \]

where \( y_t^v = \frac{V_{t+1} - \theta_x - J_{t+1} Z_{t+1}^v}{V_t} \) and \( \sigma_t^v = \frac{V_t - \theta_x}{V_t} \).

Stack up to get matrix \( X_v \) and \( Y_v \). Assume conjugate prior such that \( \beta_v \sim N(a, A) \), then the conditional posterior for \( \beta_v \) is normal:

\[ \beta_v | \sigma_v^2, \{ V_t \}_{t=1}^T \sim N(a^*, A^*) \]

where

\[ A^* = \left( \frac{(X_v)^T X_v}{\sigma_v^2} + A^{-1} \right)^{-1} \]
\[ a^* = A^* \left( \frac{(X_v)^T Y_v}{\sigma_v^2} + A^{-1} a \right) \]

**Conditional Posteriors for \( \rho \) and \( \sigma_v^2 \)**

Note that the long run mean process \( \theta_t \) is conditionally independent, the posterior can be derived as follows:

\[ p(\Theta | V, Y) \propto p(Y, V, \theta | \Theta)p(\Theta) \]
\[ \propto \prod_{t=1}^T p(y_t, V, \epsilon | V_{t-1}, \Theta)p(\Theta) \]
\[ \propto \prod_{t=1}^T |\Sigma|^{-1/2} (2\pi)^{-1/2} \exp \left( -\frac{1}{2} R_{t+1} \Sigma^{-1} R_{t+1}^T \right)p(\Theta) \]
\[ \propto |\Sigma|^{-T/2} (2\pi)^T \exp \left( -\sum_{t=1}^T \frac{1}{2} R_{t+1} \Sigma^{-1} R_{t+1}^T \right)p(\Theta) \]
For notational simplicity, we let:

\[
\sum_{t=1}^{T} R_t^T R_t = \begin{pmatrix}
  r_{11} & r_{12} & r_{13} \\
  r_{21} & r_{22} & r_{23} \\
  r_{31} & r_{32} & r_{33}
\end{pmatrix}
\]

Assume \( \psi | \Omega \sim N(c, \Omega C) \) and \( \Omega \sim IG(b, B) \), then:

\[
p(\psi | \Omega)p(\Omega) = \Omega^{-1/2} \exp\left(-\frac{1}{2} \frac{(\psi - c)^2}{\Omega C}\right) \left(\frac{1}{\Omega}\right)^{b+1} \exp\left(-\frac{B}{\Omega}\right)
\]

Hence, we are the same as the log-SV case, but with a different \( R \) matrix:

\[
\Omega | V, Y \sim IG(b^*, B^*)
\]

\[
\psi | \Omega, V, Y \sim N(c^*, \Omega C^*)
\]

where

\[
c^* = (C^*)^{-1}(r_{12} + \frac{c}{G})
\]

\[
C^* = (r_{11} + \frac{1}{G})^{-1}
\]

\[
b^* = b + \frac{T}{2}
\]

\[
B^* = B + \frac{1}{2}(r_{22} - \frac{r_{21}^2}{r_{11}})
\]

Conditional Posteriors for Volatility

\[
p(V_t | Y, V_{-t}, \theta, \Theta)
\]

\[
\propto p(V_t | y_{t+1}, y_t, V_{t-1}, V_{t+1}, \theta_{t-1}, \theta_t, \theta_{t+1}, IV_{t, r_1}, IV_{t, r_2}, IV_{t+1, r_1}, IV_{t+1, r_2}, \Theta)
\]

\[
\propto p(y_{t+1}, y_t, V_{t-1}, V_{t+1}, \theta_{t-1}, \theta_t, \theta_{t+1}, IV_{t, r_1}, IV_{t, r_2}, IV_{t+1, r_1}, IV_{t+1, r_2}, \Theta)
\]
\[ p(IV_{t+1,1}, IV_{t+1,2}, y_{t+1}, V_{t+1}, \theta_{t+1}|V_t, \Theta) p(IV_{t,1}, IV_{t,2}, y_t, V_t, \theta_t|V_{t-1}, \Theta) \]
\[ \propto \frac{1}{V_t} \exp\left(-\frac{1}{2} (R_{t+1}^{-1} R_{t+1}^T + R_t^{-1} R_t^T)\right) p(IV_{t,1}) p(IV_{t,2}) \]
\[ \propto \frac{1}{V_t} \exp\left(-\frac{1}{2} (\text{tr}(R_{t+1}^T R_{t+1} + R_t^T R_t)\Sigma^{-1})\right) p(IV_{t,1}) p(IV_{t,2}) \]

where

\[ p(IV_{t,\tau}) = \frac{1}{\sqrt{2\pi \sigma^2}} \exp\left(-\frac{1}{2\sigma^2} (IV_{t,\tau} - \alpha IV(\tau) - \beta_{IV}(\tau) V_t - \beta_{\theta IV}(\tau) \theta_t)^2\right) \]

We use the random-walk algorithm to update \( V_t \) according to the density above.

**Conditional Posteriors for Long-run Mean of Volatility**

\[ p(\theta_t|Y, V_{-t}, \theta, \Theta) \]
\[ \propto p(\theta_t|y_{t+1}, y_t, V_{t-1}, V_t, V_{t+1}, \theta_{t-1}, \theta_{t+1}, IV_{t,1}, IV_{t,2}, IV_{t+1,1}, IV_{t+1,2}, \Theta) \]
\[ \propto p(y_{t+1}, y_t, V_{t-1}, V_t, V_{t+1}, \theta_{t-1}, \theta_{t+1}, IV_{t,1}, IV_{t,2}, IV_{t+1,1}, IV_{t+1,2}, \Theta) \]
\[ \propto p(IV_{t+1,1}, IV_{t+1,2}, y_{t+1}, V_{t+1}, \theta_{t+1}|V_t, \Theta) p(IV_{t,1}, IV_{t,2}, y_t, V_t, \theta_t|V_{t-1}, \Theta) \]
\[ \propto \frac{1}{\sqrt{\theta_t}} \exp\left(-\frac{1}{2} (R_{t+1}^{-1} R_{t+1}^T + R_t^{-1} R_t^T)\right) p(IV_{t,1}) p(IV_{t,2}) \]
\[ \propto \frac{1}{\sqrt{\theta_t}} \exp\left(-\frac{1}{2} (\text{tr}(R_{t+1}^T R_{t+1} + R_t^T R_t)\Sigma^{-1})\right) p(IV_{t,1}) p(IV_{t,2}) \]

where

\[ p(IV_{t,\tau}) = \frac{1}{\sqrt{2\pi \sigma^2}} \exp\left(-\frac{1}{2\sigma^2} (IV_{t,\tau} - \alpha IV(\tau) - \beta_{IV}(\tau) V_t - \beta_{\theta IV}(\tau) \theta_t)^2\right) \]

Again, we use the random-walk algorithm to update \( \theta_t \) according to the density above.
Conditional Posteriors for $\alpha_\theta$ and $\beta_\theta$ of Long-run Mean of Volatility

Let
\[ y_t^\theta = x_t^\theta \beta + \sigma_\theta \varepsilon_{t-1} \]
where $y_t^\theta = \frac{\theta_{t+1}}{\sqrt{\theta_t}}$, $x_t^\theta = [1 \quad \theta_t]/\sqrt{\theta_t}$ and $\beta = \begin{bmatrix} \alpha_\theta & \beta_\theta \end{bmatrix}$.

Stack up to get matrix $X_\theta$ and $Y_\theta$. Assume conjugate prior such that $\alpha_\theta, \beta_\theta \sim N(a, A)$, then the conditional for $\alpha_\theta$ and $\beta_\theta$ is bivariate normal:
\[ \alpha_\theta, \beta_\theta | \sigma_\theta^2, \{\theta_t\}_{t=1}^T \sim N(a^*, A^*) \]
where
\[ A^* = \left( \frac{(X_\theta)^T X_\theta}{\sigma_\theta^2} + A^{-1} \right)^{-1} \]
\[ a^* = A^* \left( \frac{X_\theta}{\sigma_\theta^2} + A^{-1} a \right) \]

Conditional Posteriors for $\sigma_\theta^2$ of Long-run Mean of Volatility

Assume a prior $\sigma_\theta^2 \sim IG(b, B)$, then the posterior is given by:
\[ \sigma_\theta^2 | \alpha_\theta, \beta_\theta, \theta \sim IG(b^*, B^*) \]
where
\[ b^* = 0.5T + b \]
\[ B^* = 0.5 \sum_{t=1}^{T-1} \xi_t^2 + B \]
\[ \xi_t = \frac{\theta_{t+1} - \alpha_\theta - \beta_\theta \theta_t}{\sqrt{\theta_t}} \]
Jumps in Volatility: \( \mu_v \)

Assume the prior on \( \mu_v \sim IG(b, B) \), and note that conditional on the data, by Bayes's rule, the posterior is:

\[
p(\mu_v|Y, \Theta) \propto p(Z_v|Y, \Theta)p(\mu_v)
\]
\[
\propto \frac{1}{\mu_v^b} e^{-\frac{\sum_{t=1}^{n} z_{v_t}}{\mu_v}} B^b \frac{1}{\Gamma(b)} \mu_v^{b-1} e^{-\frac{B}{\mu_v}}
\]
\[
\propto \frac{1}{\mu_v^{n+b-1}} e^{-\frac{\sum_{t=1}^{n} z_{v_t}^2 + B}{\mu_v}}
\]

The posterior is hence:

\[
\mu_v|Z_v \sim IG(b^*, B^*)
\]

where

\[
b^* = n + b
\]
\[
B^* = \sum_{t=1}^{n} z_{v_t}^2 + B
\]

\( \tau_i \) is the identified jump time and there are \( n \) total number of identified jump time.

Jumps in Volatility: States update

\[
p(Z_v^{|Y, \Theta}) \propto p(y_t, V_t, \theta_t)p(Z_t)
\]
\[
\propto \exp\left(-\frac{1}{2} \text{tr}(R_t^T R_t) \Sigma^{-1} \right) \mu_v \exp\left(-\mu_v Z_v^2 \right)
\]
\[
\propto \exp\left(-\frac{1}{2} \text{tr}(R_t^T R_t) \Sigma^{-1} \right) - \mu_v Z_v^2
\]

Completing the square on \( Z_t^v \) yields the posterior to be a truncated normal:

\[
Z_v^{|Y = 1, V, \mu_v} \sim N(a^*, A^*)1_{Z_v^v > 0}
\]
where

\[ a^* = V_t - (1 - \beta_u)\theta_{t-1} - \beta_0 V_{t-1} - \sigma_v^2 (\rho^2 r_t + \mu_0 (1 - \rho^2) V_{t-1}) \]

\[ A^* = \sigma_v^2 (1 - \rho^2) V_{t-1} \]

If \( J_t = 0 \), we draw from its prior distribution, \( Z_t^\nu \sim Exp(\mu_\nu) \).
Table 1.1: Factor loadings from Principal Component Analysis on at-the-money options on S&P 500 obtained from OptionsMetric for maturities of 30, 60, 90, 122, 152, 182, 273 and 365 days. The sample period is January 4, 1996 to June 5, 2006.

<table>
<thead>
<tr>
<th>Maturity</th>
<th>30</th>
<th>60</th>
<th>90</th>
<th>122</th>
<th>152</th>
<th>182</th>
<th>273</th>
<th>365</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.4060</td>
<td>0.6506</td>
<td>-0.5947</td>
<td>-0.2317</td>
<td>0.0623</td>
<td>-0.0093</td>
<td>0.0159</td>
<td>0.0171</td>
</tr>
<tr>
<td></td>
<td>0.3894</td>
<td>0.3021</td>
<td>0.3101</td>
<td>0.5783</td>
<td>-0.5026</td>
<td>0.2483</td>
<td>-0.0727</td>
<td>-0.0841</td>
</tr>
<tr>
<td></td>
<td>0.3757</td>
<td>0.1300</td>
<td>0.3901</td>
<td>0.2413</td>
<td>0.7490</td>
<td>-0.2442</td>
<td>0.0039</td>
<td>0.1043</td>
</tr>
<tr>
<td></td>
<td>0.3551</td>
<td>-0.0624</td>
<td>0.3155</td>
<td>-0.4477</td>
<td>-0.3823</td>
<td>-0.6426</td>
<td>-0.0040</td>
<td>-0.1048</td>
</tr>
<tr>
<td></td>
<td>0.3409</td>
<td>-0.1803</td>
<td>0.1828</td>
<td>-0.3382</td>
<td>-0.0366</td>
<td>0.4759</td>
<td>0.5200</td>
<td>0.4530</td>
</tr>
<tr>
<td></td>
<td>0.3313</td>
<td>-0.2531</td>
<td>0.0411</td>
<td>-0.2739</td>
<td>0.1786</td>
<td>0.4244</td>
<td>-0.2848</td>
<td>-0.6755</td>
</tr>
<tr>
<td></td>
<td>0.3134</td>
<td>-0.3875</td>
<td>-0.2460</td>
<td>0.0712</td>
<td>-0.0546</td>
<td>-0.0064</td>
<td>-0.6499</td>
<td>0.5106</td>
</tr>
<tr>
<td></td>
<td>0.3035</td>
<td>-0.4669</td>
<td>-0.4504</td>
<td>0.3984</td>
<td>-0.0161</td>
<td>-0.2431</td>
<td>0.4697</td>
<td>-0.2201</td>
</tr>
</tbody>
</table>

Table 1.2: Variance decomposition from Principal Component Analysis on at-the-money options on S&P 500 obtained from OptionsMetric for maturities of 30, 60, 90, 122, 152, 182, 273 and 365 days. The sample period is January 4, 1996 to June 5, 2006.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Percentage Explained</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor 1</td>
<td>97.73%</td>
</tr>
<tr>
<td>Factor 2</td>
<td>2.07%</td>
</tr>
<tr>
<td>Factor 3</td>
<td>0.14%</td>
</tr>
<tr>
<td>Factor 4</td>
<td>0.03%</td>
</tr>
<tr>
<td>Factor 5</td>
<td>0.01%</td>
</tr>
<tr>
<td>Factor 6</td>
<td>0.01%</td>
</tr>
<tr>
<td>Factor 7</td>
<td>0.00%</td>
</tr>
<tr>
<td>Factor 8</td>
<td>0.00%</td>
</tr>
</tbody>
</table>

Table 1.3: Summary Statistics of S&P 500 Returns. This table provides summary statistics for daily return data on the S&P 500 for various periods described.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>January 2, 1986 - June 5, 2006</td>
<td>0.0357</td>
<td>1.0804</td>
<td>-2.0596</td>
<td>43.8481</td>
<td>-22.8997</td>
<td>43.8481</td>
</tr>
<tr>
<td>November 2, 1987 - June 5, 2006</td>
<td>0.0352</td>
<td>1.0114</td>
<td>-0.2353</td>
<td>4.2731</td>
<td>-7.1127</td>
<td>5.5744</td>
</tr>
</tbody>
</table>
Table 1.4: Summary Statistics of VIX Index. This table provides summary statistics for daily VIX data for various periods described.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>Skew</th>
<th>Kurt.</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>January 2, 1986 - June 5, 2006</td>
<td>0.2074</td>
<td>0.0817</td>
<td>3.1429</td>
<td>30.6630</td>
<td>0.0904</td>
<td>1.5019</td>
</tr>
<tr>
<td>November 2, 1987 - June 5, 2006</td>
<td>0.2052</td>
<td>0.0760</td>
<td>1.1270</td>
<td>1.5447</td>
<td>0.0904</td>
<td>0.5847</td>
</tr>
</tbody>
</table>

Table 1.5: Summary Statistics of long-dated option implied volatilities. This table provides summary statistics for daily VIX data for options of various maturity for the sample from January 4, 1996 to June 5, 2006.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>Skew</th>
<th>Kurt.</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>30 Days IV</td>
<td>1.5339</td>
<td>1.0083</td>
<td>1.6879</td>
<td>3.8815</td>
<td>0.2628</td>
<td>7.2811</td>
</tr>
<tr>
<td>60 Days IV</td>
<td>1.5506</td>
<td>0.9297</td>
<td>1.4561</td>
<td>2.9911</td>
<td>0.3394</td>
<td>6.8128</td>
</tr>
<tr>
<td>90 Days IV</td>
<td>1.5540</td>
<td>0.8866</td>
<td>1.3934</td>
<td>3.0122</td>
<td>0.3899</td>
<td>6.7635</td>
</tr>
<tr>
<td>122 Days IV</td>
<td>1.5499</td>
<td>0.8258</td>
<td>1.2541</td>
<td>2.6329</td>
<td>0.4024</td>
<td>6.3871</td>
</tr>
<tr>
<td>152 Days IV</td>
<td>1.5495</td>
<td>0.7896</td>
<td>1.1830</td>
<td>2.5663</td>
<td>0.4201</td>
<td>6.3242</td>
</tr>
<tr>
<td>182 Days IV</td>
<td>1.5499</td>
<td>0.7670</td>
<td>1.1325</td>
<td>2.4589</td>
<td>0.4384</td>
<td>6.1631</td>
</tr>
<tr>
<td>273 Days IV</td>
<td>1.5543</td>
<td>0.7272</td>
<td>0.9719</td>
<td>1.8159</td>
<td>0.4656</td>
<td>5.6955</td>
</tr>
<tr>
<td>365 Days IV</td>
<td>1.5643</td>
<td>0.7097</td>
<td>0.8801</td>
<td>1.4908</td>
<td>0.4925</td>
<td>5.5349</td>
</tr>
</tbody>
</table>

Figure 1.1: Term-structure of Implied Volatility on representative days and the average term-structure implied volatility over the sample 1996 to 2006.
<table>
<thead>
<tr>
<th></th>
<th>SV</th>
<th>SVRJ</th>
<th>SVCJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\kappa$</td>
<td>1.1435 (0.0410)</td>
<td>1.1439 (0.0375)</td>
<td>2.1953 (0.0921)</td>
</tr>
<tr>
<td>$\theta$</td>
<td>0.0375 (0.0002)</td>
<td>0.0376 (0.0002)</td>
<td>0.0207 (0.0009)</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.0913 (0.0017)</td>
<td>0.0930 (0.0016)</td>
<td>0.0849 (0.0017)</td>
</tr>
<tr>
<td>$\rho$</td>
<td>-0.7562 (0.0103)</td>
<td>-0.7904 (0.0107)</td>
<td>-0.7868 (0.0106)</td>
</tr>
<tr>
<td>$\mu_2$</td>
<td>0.3098 (0.6301)</td>
<td>-1.7317 (0.6018)</td>
<td></td>
</tr>
<tr>
<td>$\sigma_2$</td>
<td>3.4729 (0.7373)</td>
<td>2.9450 (0.6051)</td>
<td></td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.0095 (0.0031)</td>
<td>0.0074 (0.0017)</td>
<td></td>
</tr>
<tr>
<td>$\mu$</td>
<td>0.0286 (0.0013)</td>
<td>0.0266 (0.0029)</td>
<td>0.0362 (0.0022)</td>
</tr>
<tr>
<td>$\mu_v$</td>
<td>0.8851 (0.1481)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_{\gamma_1}$</td>
<td>0.3025 (0.0042)</td>
<td>0.2975 (0.0040)</td>
<td>0.2978 (0.0042)</td>
</tr>
<tr>
<td>$\alpha_{\gamma_1}$</td>
<td>0.3432 (0.0023)</td>
<td>0.3302 (0.0004)</td>
<td>0.3444 (0.0014)</td>
</tr>
<tr>
<td>$\beta_{\gamma_1}^v$</td>
<td>0.9995 (0.0003)</td>
<td>0.9996 (0.0004)</td>
<td>0.9998 (0.0003)</td>
</tr>
<tr>
<td>$\sigma_{\gamma_2}$</td>
<td>0.0224 (0.0007)</td>
<td>0.0221 (0.0007)</td>
<td>0.0232 (0.0008)</td>
</tr>
<tr>
<td>$\alpha_{\gamma_2}$</td>
<td>0.4537 (0.0038)</td>
<td>0.4510 (0.0033)</td>
<td>0.4526 (0.0032)</td>
</tr>
<tr>
<td>$\beta_{\gamma_2}^v$</td>
<td>0.6121 (0.0048)</td>
<td>0.6088 (0.0043)</td>
<td>0.6138 (0.0048)</td>
</tr>
</tbody>
</table>

Table 1.6: S&P 500 Parameter Estimates of One-factor Volatility Model with Two Options for the sample period 1986-2006. Parameter estimates for the two-factor volatility model using returns and VIX data for the sample January 2, 1986 to June 5, 2006, and the one-year option from January 2, 1996 to June 5, 2006. The models and parameterizations are given in Section II and the estimates correspond to percentage changes in the index value.

Figure 1.2: Residuals of returns equation from fitting the two-factor. The red, blue and black line represent the SV, SVJ and SVCJ model correspondingly.
### Table 1.7: S&P 500 Parameter Estimates of Two-factor Volatility Model with Two Options for the sample period 1986-2006

<table>
<thead>
<tr>
<th></th>
<th>SV</th>
<th>SVRJ</th>
<th>SVCJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\kappa_v$</td>
<td>2.9666 (0.5312)</td>
<td>2.1253 (0.4152)</td>
<td>5.4422 (0.6374)</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.1511 (0.0022)</td>
<td>0.1570 (0.0023)</td>
<td>0.1440 (0.0021)</td>
</tr>
<tr>
<td>$\kappa_\theta$</td>
<td>0.6061 (0.0419)</td>
<td>0.5604 (0.0441)</td>
<td>0.5703 (0.0436)</td>
</tr>
<tr>
<td>$\theta_\theta$</td>
<td>0.0417 (0.0007)</td>
<td>0.0493 (0.0015)</td>
<td>0.0338 (0.0004)</td>
</tr>
<tr>
<td>$\sigma_\theta$</td>
<td>0.0520 (0.0019)</td>
<td>0.0580 (0.0020)</td>
<td>0.0476 (0.0017)</td>
</tr>
<tr>
<td>$\rho$</td>
<td>-0.7470 (0.0075)</td>
<td>-0.8046 (0.0090)</td>
<td>-0.7600 (0.0081)</td>
</tr>
<tr>
<td>$\mu_z$</td>
<td>0.7149 (0.1721)</td>
<td>-1.2592 (0.4391)</td>
<td></td>
</tr>
<tr>
<td>$\sigma_z$</td>
<td>1.7645 (0.2035)</td>
<td>3.3503 (0.4026)</td>
<td></td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.0378 (0.0077)</td>
<td>0.0122 (0.0021)</td>
<td></td>
</tr>
<tr>
<td>$\mu$</td>
<td>0.0271 (0.0012)</td>
<td>-0.0050 (0.0045)</td>
<td>0.0329 (0.0025)</td>
</tr>
<tr>
<td>$\nu$</td>
<td>1.0432 (0.1409)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_{r_1}$</td>
<td>0.0706 (0.0019)</td>
<td>0.0698 (0.0019)</td>
<td>0.0692 (0.0018)</td>
</tr>
<tr>
<td>$\alpha_{r_1}$</td>
<td>0.2205 (0.0061)</td>
<td>0.2368 (0.0066)</td>
<td>0.2168 (0.0056)</td>
</tr>
<tr>
<td>$\beta_{r_1}^a$</td>
<td>0.9994 (0.0008)</td>
<td>0.9994 (0.0009)</td>
<td>0.9994 (0.0008)</td>
</tr>
<tr>
<td>$\beta_{r_1}^g$</td>
<td>0.1435 (0.0062)</td>
<td>0.1304 (0.0070)</td>
<td>0.2182 (0.0069)</td>
</tr>
<tr>
<td>$\sigma_{r_2}$</td>
<td>0.0397 (0.0011)</td>
<td>0.0409 (0.0011)</td>
<td>0.0415 (0.0011)</td>
</tr>
<tr>
<td>$\alpha_{r_2}$</td>
<td>0.3984 (0.0064)</td>
<td>0.4208 (0.0071)</td>
<td>0.4027 (0.0107)</td>
</tr>
<tr>
<td>$\beta_{r_2}^a$</td>
<td>0.2727 (0.0066)</td>
<td>0.2708 (0.0056)</td>
<td>0.2684 (0.0061)</td>
</tr>
<tr>
<td>$\beta_{r_2}^g$</td>
<td>0.4377 (0.0082)</td>
<td>0.3644 (0.0088)</td>
<td>0.5497 (0.0100)</td>
</tr>
</tbody>
</table>

Table 1.7: S&P 500 Parameter Estimates of Two-factor Volatility Model with Two Options for the sample period 1986-2006. Parameter estimates for the two-factor volatility model using returns and VIX data for the sample January 2, 1986 to June 5, 2006, and the one-year option from January 2, 1996 to June 5, 2006. The models and parameterizations are given in Section II and the estimates correspond to percentage changes in the index value.

### Table 1.8: In-sample pricing error for the sample period 1986-2006

<table>
<thead>
<tr>
<th></th>
<th>SV</th>
<th>SVCJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>One-factor (30 days)</td>
<td>0.0376</td>
<td>0.0338</td>
</tr>
<tr>
<td>One-factor (365 days)</td>
<td>0.0021</td>
<td>0.0022</td>
</tr>
<tr>
<td>Two-factor (30 days)</td>
<td>0.0248</td>
<td>0.017</td>
</tr>
<tr>
<td>Two-factor (365 days)</td>
<td>0.0035</td>
<td>0.0036</td>
</tr>
</tbody>
</table>

Table 1.8: In-sample pricing error for the sample period 1986-2006. Parameter estimates for the two-factor volatility model using returns and VIX data for the sample January 2, 1986 to June 5, 2006, and the one-year option from January 2, 1996 to June 5, 2006. Since the two-factor models are estimated from both the 30 days and 365 days at-the-model options, their relevant pricing errors are both computed.

<table>
<thead>
<tr>
<th></th>
<th>One-factor</th>
<th>Two-factor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SV</td>
<td>SVCJ</td>
</tr>
<tr>
<td>( \kappa_C )</td>
<td>1.0774 (0.0191)</td>
<td>1.0709 (0.0190)</td>
</tr>
<tr>
<td>( \kappa_S )</td>
<td>0.0015 (0.0003)</td>
<td>1.3971 (0.0517)</td>
</tr>
<tr>
<td>( \lambda )</td>
<td>-2.4308 (0.0000)</td>
<td>-1.2682 (0.4410)</td>
</tr>
<tr>
<td>( \mu_C )</td>
<td>2.9659 (0.6174)</td>
<td>3.3689 (0.4171)</td>
</tr>
<tr>
<td>( \sigma_C )</td>
<td>0.0793 (0.0478)</td>
<td>0.0481 (0.0390)</td>
</tr>
</tbody>
</table>

Table 1.10: Out-of-sample pricing error for the sample period 1996-2006. Using the results estimated using VIX, where maturity is 30 days, from 1987-2006, the table displays pricing error in RMSE for the options of maturities not used in the estimation for the one-factor model. The pricing error are computed in implied volatility units.

<table>
<thead>
<tr>
<th></th>
<th>One-factor</th>
<th>Two-factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maturity</td>
<td>SV</td>
<td>SVCJ</td>
</tr>
<tr>
<td>60</td>
<td>0.0273</td>
<td>0.0268</td>
</tr>
<tr>
<td>90</td>
<td>0.0235</td>
<td>0.0227</td>
</tr>
<tr>
<td>122</td>
<td>0.0201</td>
<td>0.0188</td>
</tr>
<tr>
<td>152</td>
<td>0.0177</td>
<td>0.016</td>
</tr>
<tr>
<td>182</td>
<td>0.0159</td>
<td>0.0136</td>
</tr>
<tr>
<td>273</td>
<td>0.0121</td>
<td>0.008</td>
</tr>
<tr>
<td>Average</td>
<td>0.0194</td>
<td>0.0177</td>
</tr>
</tbody>
</table>
Table 1.11: This table provides a summary of the simulation results for the SVCJ model. We sample a total of 100,000 samples and burn out 50,000 to get our final estimates.
Figure 1.3: Time-series Estimates of latent processes of the one-factor SVCJ model using MCMC. The estimation uses joint time-series of S&P 500 returns, VIX index and the one-year option on the index. The sample period is January 2, 1987-June 5, 2006. The models and parameterizations are given in Section II. The top panel displays stochastic volatility (thin black line) and long-run mean of volatility (thick red line). The medium shows jump sizes in returns and the bottom panel shows jump sizes in volatility.
Figure 1.4: Time-series Estimates of latent processes of the two-factor SVCJ model using MCMC. The estimation uses joint time-series of S&P 500 returns, VIX index and the one-year option on the index. The sample period is January 2, 1987-June 5, 2006. The models and parameterizations are given in Section II. The top panel displays stochastic volatility (thin black line) and long-run mean of volatility (thick red line). The medium shows jump sizes in returns and the bottom panel shows jump sizes in volatility.
Figure 1.5: Time-series Estimates of latent processes for the simulation study using returns only. The four panels correspond to estimates of volatility, long-run mean, jump sizes in returns and jump sizes in volatility respectively. The black line displays the true process, and the red line corresponds to the MCMC estimates.
Figure 1.6: Time-series Estimates of latent processes for the simulation study using jointly returns and one short-dated option. The four panels correspond to estimates of volatility, long-run mean, jump sizes in returns and jump sizes in volatility respectively. The black line displays the true process, and the red line corresponds to the MCMC estimates.
Figure 1.7: Time-series Estimates of latent processes for the simulation study using jointly returns, one short-dated option and one long-dated option. The four panels correspond to estimates of volatility, long-run mean, jump sizes in returns and jump sizes in volatility respectively. The black line displays the true process, and the red line corresponds to the MCMC estimates.
Chapter 2

Shocks to Asset Pricing Factors: Understanding the Risks and Realized Returns of Common Asset Pricing Factors

2.1 Introduction

Since Fama and French (1993), observable, quantitative, cross-sectional, asset pricing factors, such as those generated by sorting stocks based on size (SMB), metric of “value” such as market to book (HML), or past price performance (momentum/UMD), play an important role in both academic research and portfolio management. The evidence supporting these common and others factors is quite strong, although the underlying economic sources of the common variation and factor premia are not well understood. In particular, there is still a vigorous debate regarding the source, factor premia or mispricing, of the historical excess performance of some of these factors such as value or momentum.

Despite the large amount of research focused on these factors, little is known
about the structure of the underlying shocks driving these factors. Recent market experiences in 2007 and 2008 suggest that the time series returns generated by these factors are quite complicated, with strong time-variation in volatility and potentially extremely large outliers. These dynamics are reminiscent of aggregate equity index returns, where there are well known outlier events such as the stock market crash of 1987 and index volatility varies strongly over time.

The goal of this paper is to fill this gap in the literature. We provide stylized facts about the shocks driving common asset pricing factors such as size, value, and momentum using flexible stochastic volatility (SV) models that are widely used for modeling index returns. These models decompose returns into three components: a predictable mean component capturing average returns as a function of volatility, a normally distributed shock term with time-varying volatility, and a rare jump or outlier component that captures periods of exceptional stress.

Volatility is mean-reverting and its shocks are similar in form to those in returns. The normally distributed shocks in volatility are correlated with the corresponding shocks in returns (the leverage effect) as are the large, positive jump-type shocks that occurs coincident to the large movements in returns. This latter component captures periods of market stress in which volatility tends to dramatically increase at the same time returns prices fall. The model provides a very rich specification for the risks underlying these factors. In analyzing these factors, much of the literature focuses on unconditional moments: average returns and Sharpe ratios, neglecting the non-normal

---

1 Examples of these movement abounded in the mainstream press and research reports. For example, Goldman, Sachs, CFO reported that “We were seeing things that were 25-standard deviation moves, several days in a row” (Larsen 2007). Another analyst, referring to quantitative asset pricing factors, stated the following: “As for what has been transpiring in August, we have been able to document daily returns of this magnitude occurring before only at the height of the bursting of the Internet Bubble and in the late 1960s. This appears to be an event with little precedent” (Rothman (2007)). On January 23, 2008, the “momentum” factor had one its worst days ever: “Price momentum experienced its fourth-biggest loss ever (measured since July 1950) on Wednesday as beaten down stocks soared and prior winners plunged. To put Wednesday’s loss in context, there have only been six days since July 1950 with greater under-performance by momentum, all of which were in 2001 as the Internet bubble was popping” (Rothman (2008)).

2 These outliers movements are related to buy distinct from factor model “misbehavior.”
risk components and variation of these risks over time.

To estimate the models, we use relatively high-frequency daily returns from 1980 to 2007 for the four factors mentioned above. Although the literature typically uses monthly or even quarterly data, daily data offers a number of advantages. On the statistical side, higher frequency data provides more accurate volatility estimates, and importantly increases the ability to disentangle jump from normal components. When prices are sampled at lower frequencies, distinctions between different shocks are blurred by time-aggregation. On the practical side, the size, value, and momentum portfolios are all leveraged and for asset managers, these high frequency fluctuations cannot be ignored, as margin and collateral must be moved at a daily frequency.

For statistical inference, we take a Bayesian perspective. The Bayesian approach, implemented via MCMC, is commonly used in this class of models as it provides a computational feasible approach to parameter and latent state variable estimation, as it is easy to estimate latent volatility, jump times, and jump sizes using the output of the MCMC algorithm. The Bayesian approach also provides a formal mechanism for imposing prior information, which is often required in mixture models (Kiefer (1978)). These identification problems can be formally treated via the prior distribution. In the case of the models under consideration here, to ameliorate the identification problem, one needs to assume that jump movements are large, a natural assumption when dealing with outliers.

Conditional on parameter estimates, we use the particle filter to compute predictive likelihoods for likelihood ratio statistics. Unlike most of the literature, we focus on sequential likelihood ratios, tracking the relative performance of different models over time. In contrast to full-sample tests (which we also report), this sequential model monitoring provides a unique view of how competing models fit the data on a day-to-day basis. For example, it identifies exact days or periods of time in which one model outperforms another.

Turning to the empirical results, we first discuss the results for the market port-
folio. Although similar to previous results in the literature, they are useful to summarize. Unconditionally, market returns have strong excess kurtosis and negative skewness. First, there is a strong negative leverage effect for both normal distributed shocks to volatility and jumps. When volatility increases via a normally distributed shocks, returns tend to decrease, with a correlation of about -50%. The leverage effect is also there in the extremes. Jumps arrive at a rate of slightly more than 1 per year, resulting in a negative shock to returns of about -1.5% to -2.5% and a positive shock that increases volatility from about 15% to 23%, on average. Second, market volatility is quite volatile. One way to see this is to use the estimated volatility states and compare the ratio of the 95th to 5th quantiles of the estimated paths, which is about 3.4 for the market. This high and variable volatility is a difficult challenge for asset pricing models. Third, there is no systematic relationship between average returns and volatility. The relationship is unstable across models and statistically insignificant.

Overall, market returns look very much like a “carry” trade. In normal times (periods without jumps), the market delivers very high average returns, about 10% to 11% on an annualized basis. These high returns in part reflect compensation for jump risk: in periods of market stress, when jumps occur, returns fall dramatically and volatility increases dramatically. Over long periods of time, the normal and market stress periods average out delivering the historical 7.5% average market return. A comparison of the likelihood ratios overtime strongly supports the models with jumps in both returns and volatility.

Turning next to factor portfolios, the most interesting results are for the value and momentum portfolios. For the value factor, the first interesting observation is that it has excess kurtosis but is positively skewed, an uncommon feature for equity returns. In terms of leverage effects, both normal and market stress leverage effects are positive, or at least non-negative. When volatility increases via a normal shock, returns increase slightly on average, although the effect is not significant. When
volatility increases dramatically via a jump, the mean jump size in returns is slightly positive, and in the most general specification, is greater than 1%. The volatility jumps are quite substantial, more than doubling HML's volatility from an average of 7.5% to over 16%. Finally, a decomposition of HML's average returns indicates that HML's average returns increases with volatility. The conditional mean return during normal periods is $\alpha + \beta V_t$, and for HML $\alpha$ is close to zero and insignificant, but $\beta$ is large and statistically significant across all specifications.

Based on this, we obtain a very different view of HML, compared to the market portfolio. Unlike the market which performs poorly during periods of stress (when its volatility is high or increasing), HML performs well in periods of market stress. As volatility increases, via a normally distributed or jump shock, HML has positive returns on average via the two leverage effects. Moreover, when volatility is high, HML's average return is also high. In this regard, HML has every feature of statistical "arbitrage": high unconditional returns, positive skewness, positive returns as volatility increases, and high average returns when volatility is high. The sequential likelihood ratios vary substantially over the sample, but no model clearly dominates, unlike market returns.

The return decompositions for the momentum factor are almost the polar opposite of HML. For momentum, $\beta$ is negative and highly significant, while $\alpha$ is very large, on the order of 15%, in annualized units. When volatility increases via a jump, which arrive at a rate of about 1 per year, momentum has significantly negative returns, on the order of -1% to -2% and annualized volatility increases from about 8% to 16%. Interestingly, the normal leverage effect for momentum is modestly positive and significant, indicating that normal shock in momentum's volatility leads to positive average returns. In contrast to the value factor, momentum performs poorly in stressed states and is also classic "carry" asset: when volatility is low, momentum has high average returns, but when volatility increases dramatically or is high, UMD's average returns are very low.
The time series of volatilities is also quite interesting. Both momentum and HML have highly variable volatility, as the ratio 95th to 5th quantiles of the estimated volatility paths is 5.3 for momentum and 4.2. For example, there are years where HML's conditional volatility is as low as 6% and as high as 35%, according to the estimates. For momentum, the variation is similar, moving regularly from very low rates of 3% to 4% to more than 20%. Thus, the risks underlying these factors varies dramatically over time, maybe even more so than market volatility. Given our inability to explain the sources of high frequency time-variation in aggregate market indices, these results indicate that these asset prices have similar high frequency variation that standard low-frequency models will have difficulties explaining.

The estimated volatilities also have interesting comovements. Over the full sample, the factor volatilities are significantly correlated in levels, between 57% and 74%. However, the changes in volatilities are less correlated, and in some cases display virtually no correlation. This low correlation in changes is due to the fact that during certain periods of market stress, the factors are highly correlated, while in other periods the correlation is low. Two examples from 2007 provide the necessary intuition. After the mini-crash on February 27, 2007, market volatility increased from 6% to 15%, but SMB, HML, and UMD volatility did not increase. In July/August 2007, the volatility of value, momentum, and the market factor all roughly doubled in a few weeks.

The results in the paper provide a new set of stylized facts about common asset pricing factors. In particular, two results are particularly striking. First, the fact that the volatility of HML and UMD are so variable over time and at high frequencies. Second, the very strong relationships between HML and UMD average returns and their volatility. These results pose challenges to theories that attempt to explain the returns on these portfolios. In addition to explaining the high average returns and low volatility, these theories must also address the strong, high frequency variation in the risks in the portfolios, and the state-dependent payoff patterns of these portfolios.
The rest of the paper is outlined as follows. Section 2 introduces the models and discusses econometric issues. Section 3 provides empirical results, and Section 4 concludes.

2.2 Time series models for return decompositions

We use flexible time series models incorporating time series predictability, stochastic volatility, and "jump" or outliers movements, both in returns and in volatility. Discretely compounded returns are given by

\[ r_{t+1} = \alpha + \beta V_t + \sqrt{V_t} \varepsilon_{t+1} + J_{t+1} Z_{t+1}, \]

where \( \sqrt{V_t} \) is the stochastic volatility (to be specified below), \( \text{Prob}[J_{t+1} = 1] = \lambda \), \( \varepsilon_{t+1} \sim \mathcal{N}(0, 1) \), and \( Z_{t+1} \sim \mathcal{N}(\mu_z, \sigma_z^2) \). Here, we specify that the conditional mean contains a term depending on volatility, in contrast to existing work that analyzes time-variation in asset pricing average returns related to predictors such as the valuation ratios, dividend-yields or business cycle related variables (e.g., Rytchkov (2008)). Jumps are not predictable, so they are not related to expected return variation.

This model has multiple factors driving returns, and thus the conditional moments contain both the "normal" and "abnormal" or jump components. The mean and variance, conditional on \( V_t \) is

\[
E_t(r_{t+1}) = \alpha + \beta V_t + \lambda \mu_z \\
var_t(r_{t+1}) = V_t + \lambda (\mu_z^2 + \sigma_z^2) - \lambda^2 \mu_z^2.
\]

When \( \lambda \) is small (i.e. jumps are rare), \( \var_t(r_{t+1}) \approx V_t + \lambda (\mu_z^2 + \sigma_z^2) \). Since the asset pricing factors under consideration here are either excess market returns or long-short portfolios, there is no interest rate term in the conditional mean. At this stage, we do not interpret \( \alpha \) or \( \beta \) explicitly as prices of risks for obvious reasons. For aggregate
index returns, it is common to interpret a positive $\beta$ as a premium related to volatility. For the asset pricing factors, it is difficult to justify that the factor's own volatility is the relevant priced risk factor, although it is clear that it is a measure of risk. At a minimum, estimates of $\beta$ that are significantly positive or negative based on historical data indicate that historical returns have been high or low during periods of high volatility.

The specification for volatility is a time-discretization of Duffie, Pan, and Singleton's (2000) continuous-time square-root jump-diffusion specification:

$$V_{t+1} = \alpha_v + \beta_v V_t + \sigma_v \sqrt{V_t} \epsilon_{t+1}^v + J_{t+1} Z_{t+1}^v,$$

where the diffusive shocks are correlated, $\text{corr}(\epsilon_{t+1}, \epsilon_{t+1}^v) = \rho$, the timing of jumps in variance is coincident to the timing of jumps in returns, and the sizes of variance jumps are positive, $Z_{t+1}^v \sim \exp(\mu_v)$, to capture the fact that uncertainty rapidly increases, but slowly mean-reverts. The unconditional average variance is given by

$$E(V_t) = \frac{\alpha_v + \lambda \mu_v}{1 - \beta_v},$$

which will be used later. This model has received attention in the option pricing literature as an accurate model of the dynamics of equity index volatility, both using historical returns and option prices.

From this, we can decompose realized returns into a predictable component, $\alpha + \beta V_t$, a normally distributed $\sqrt{V_t} \epsilon_{t+1}$ that is correlated with shocks in volatility, and a rare event component, $J_{t+1} Z_{t+1}$, that is also correlated in extremes with volatility. As an example, common parameter estimates for aggregate indices imply that $\alpha > 0$, $\mu_z < 0$, $\beta$ is insignificant, and $\rho < 0$. This implies that market trends up via $\alpha$. When volatility increases locally, via a Brownian shock, stock prices tend to fall via $\rho$. When there is a jump, $J_{t+1} = 1$, volatility increases dramatically at the same time prices fall. There is no significant relationship between returns and volatility, meaning that
returns do not tend to be higher or lower when volatility is high.

2.2.1 Parameter and state estimation

Parameter estimation in this class of models is now standard using the Bayesian approach. The Bayesian approach adheres to one principle and one rule. The principle is that anything that is not directly observed should be treated as a random variable, anything observed is not a random variable. For example, this implies that parameters and unobserved states or shocks (e.g., volatility, jump times, or jump sizes) are random variables. The rule is that when making probabilistic statements, one must follow the rules of the probability, in particular Bayes rule. Bayes rule is a minimal constraint on rationality that governs updating after data is received.

For the specifications under consideration, the posterior distribution is generally of the form $p(\theta, x_T^T | r^T)$, where $\theta$ is a vector of structural parameters, $x_T^T = (x_1, ..., x_T)$, where each $x_t$ is a set of latent variables at time $t$, and $r^T$, similarly defined, is a vector of observed data up to time $T$. For the most general models under consideration, $x_t$ will contain spot variance, $V_t$, jump times, $J_t$, return jump sizes, $Z_t$, and variance jump sizes, $Z^v_t$. Notice the dimensionality of the posterior: it is a $T + K$ dimensional distribution. Because of this, characterizing the posterior using analytical methods or importance sampling methods is not possible, due to the curse of dimensionality.

Markov Chain Monte Carlo (MCMC) is the standard method for simulating posterior distributions in latent state variable models. The basic idea of the approach is simple: simulate a Markov Chain defined over $(\theta, x^T)$, where the transition kernel, $\pi[(\theta, x^T)^g, (\theta, x^T)^{(g+1)}]$, is designed so that the Markov Chain has an equilibrium distribution given by $p(\theta, x^T | r^T)$. There is wide flexibility in designing the Markov Chain, as many different chains have the same equilibrium distribution. The key to designing the chains is the Clifford-Hammersley theorem which states an appropriately chosen set of conditional distributions contains the same information as a joint distribution. In our setting, a simple application of the theorem would im-
ply that \( p(\theta|x^T, r^T) \) and \( p(x^T|\theta, r^T) \) completely characterize the joint distribution \( p(\theta, x^T|r^T) \). If it were possible to simulate from \( p(\theta|x^T, r^T) \) and \( p(x^T|\theta, r^T) \), iteratively sampling from these distributions would generate a Markov Chain with the appropriate equilibrium distribution. However, in the models that we consider, it is not possible to directly sample from these distribution.

Details of our algorithm are given in Appendix A, which provide the exact conditional distributions used to generate the Markov Chain. The prior distributions are also given. For all of the parameters, the priors are "uninformative," with the exception of \( \sigma^2_x \) and \( \mu_v \). For \( \sigma^2_x \), we calibrate the prior to have mean and standard deviation equal to three times the daily standard deviation, to insure that jumps are large movements. Without the constraint, the model will try to fit many small jumps, as noted by Kiefer (1979) and others.

### 2.2.2 Model comparison

We also provide diagnostics for model comparisons using marginal likelihoods. Jeffreys (1939) introduced the Bayesian approach to model comparison. Formally, assume there are \( \{M_k\}_{k=1}^K \) models under consideration, and the list is exhaustive. The Bayesian approach to model testing is quite simple and merely compares the posterior probability of model \( i \) to model \( j \). The Odds Ratio provides the metric and is the posterior ratio of model probabilities of model \( j \) and \( k \):

\[
\text{odds} (M_j \text{ vs. } M_k|x^t) = \frac{p(M_j|x^t)}{p(M_k|x^t)},
\]

where by Bayes rule

\[
p (M_j|x^t) = \frac{p(x^t|M_j) p(M_j)}{p(x^t)}.
\]

Thus, the Odds ratio between model \( j \) and \( k \) is

\[
\text{odds} (M_j \text{ vs. } M_k|x^t) = \frac{p(x^t|M_j) p(M_j)}{p(x^t|M_k) p(M_k)} = \mathcal{LR}_{j,k} (x^t) \cdot \text{odds} (M_j \text{ vs. } M_k),
\]
where \( \text{odds}(M_j \text{ vs. } M_k) \) is the prior odds, \( p(M_j)/p(M_k) \), and \( \mathcal{L}R_{j,k}(t^t) \) is the Bayes factor or likelihood ratio between the models:

\[
\mathcal{L}R_{j,k}(t^t) = \frac{p(r^t|M_j)}{p(r^t|M_k)}.
\]

When analyzing models over time, it is possible to recursively define the Bayes factor as

\[
\mathcal{L}R_{j,k}(t^{t+1}) = \frac{p(r_{t+1}|r^t, M_j)}{p(r_{t+1}|r^t, M_k)} \mathcal{L}R_{j,k}(t^t),
\]

where the change in the Bayes factor is driven by the new information embedded in the predictive likelihoods, \( p(r_{t+1}|r^t, M_j) \) and \( p(r_{t+1}|r^t, M_k) \).

Most applications use full-sample likelihood ratios, \( \mathcal{L}R_{j,k}(r^T) \), providing an overall measure of model fit for entire sample. The sequential, recursive formulations that we utilize are not common, but provide a much deeper understanding of how models differ. In particular, they allow researchers to discriminate between abrupt failures and those that accumulate slowly, which is impossible using full sample likelihood ratios. As seen below, it is not uncommon for one model to strongly outperform another for a long period of time and then a few observations occur that reverse the relative performance. Another interesting case is when \( \mathcal{L}R_{j,k}^T >> 1 \), but the differences are driven by a few influential observations. Ignoring these observations, the fit could be poor with, for example, \( \mathcal{L}R_{j,k}^T << 1 \) for most of the sample. This insidious case would not be diagnosed using full sample likelihoods ratios.

In settings with two models (or hypothesis), the Bayes factor is related to model probabilities. For example, when considering the probability of A against the probability of all alternatives,

\[
\text{Prob}[A] = \frac{BF(A)}{1 + BF(A)},
\]

which is just the usual comparison between odds and probabilities. For example, if the odds on A are 10:1 (BF=10), then the probability is

\[
\text{Prob}[A] = \frac{10}{1 + 10} = \frac{10}{11}.
\]

---

\(^3\) In settings with two models (or hypothesis), the Bayes factor is related to model probabilities. For example, when considering the probability of A against the probability of all alternatives,

\[
\text{Prob}[A] = \frac{BF(A)}{1 + BF(A)},
\]

which is just the usual comparison between odds and probabilities. For example, if the odds on A are 10:1 (BF=10), then the probability is

\[
\text{Prob}[A] = \frac{10}{1 + 10} = \frac{10}{11}.
\]
The predictive likelihoods used to compute the likelihood ratios are given by

\[
p(r_{t+1}|r^t, M_j) = \int p(x_{t+1}|x_t, \theta, M_j) p(x_t, \theta|r^t, M_j) dx_t, \quad (2.1)
\]

which integrates out all of the parameter and state variable uncertainty. Unfortunately, it is computationally infeasible to compute these predictive likelihoods for these models, as it requires re-running MCMC for each data point. As an alternative, we compute

\[
p(r_{t+1}|\hat{\theta}, r^t, M_j) = \int p(x_{t+1}|x_t, \hat{\theta}, M_j) p(x_t|\hat{\theta}, r^t, M_j) dx_t, \quad (2.2)
\]

where \( \hat{\theta} \) is an estimate of the parameters and \( p(x_t|\hat{\theta}, r^t, M_j) \) is the filtering distribution of the states at time \( t \). Since \( \hat{\theta} \) is not the MLE, the approximations are not directly comparable to common approximations such as AIC and BIC.

The filtering distribution, \( p\left(x_t|\hat{\theta}, r^t, M_j\right) \), where \( x_t = (J_t, Z_t, V_t) \), can be accurately computed using the particle filter. The particle filter approximates the continuous distribution, \( p\left(x_t|\hat{\theta}, r^t, M_j\right) \), by discrete distribution, \( p_N\left(x_t|\hat{\theta}, r^t, M_j\right) \), that consists of weights \( \left\{ \pi_t^{(i)} \right\}_{i=1}^N \) and support points or particles, \( \left\{ x_t^{(i)} \right\}_{i=1}^N \). A particle filtering algorithm is just a recursive algorithm for propagating the particles and computing the weights. Johannes, Polson, and Stroud (2008) provide a detailed analysis of these filtering problems for the class of models that we consider. Johannes and Polson (2008) provide a general overview of particle filtering algorithms.

2.3 Data and empirical results

2.3.1 Summary statistics and initial observations

This paper uses daily returns for four common factors: the value-weighted CRSP excess market returns, SMB, HML, and momentum. SMB and HML are constructed from six Fama-French portfolio sorted on size and book-to-market ratios. SMB, com-
monly called the size factor, is the difference in returns between the three small size portfolios (high, medium, and low book-to-market ratios) and the three large size portfolios. Thus, it is long small stocks and short large stocks. HML, commonly called the value factor, is the difference in returns between the two high book-to-market portfolios (small and large size ratios) and the two low book-to-market portfolios. Thus, HML is long high book-to-market stocks and short low book-to-market stocks. Momentum is long small and big stocks with high prior performance over the past year and short small and big stocks with low prior performance over the past year.

Table 2.1 reports summary statistics for each factor. For raw returns, the table provides annualized sample mean, daily and annualized standard deviations, skewness and excess kurtosis for each factor. Additionally, we report skewness and excess kurtosis for standardized returns, \( r_t = r_t / s_{t-1} \) where

\[
s_t^2 = \sum_{i=0}^{22} (r_{t-i} - \overline{r}_{t,22})^2
\]

is the sample variance computed using the previous 22 daily observations and \( \overline{r}_{t,22} \) is the sample mean of those prior observations. The standardized returns provide a crude correction for time-varying volatility. For normally distributed data, the skewness and excess kurtosis are both 0. Approximate standard errors for the skewness and kurtosis statistics are given by \( \sqrt{6/T} \) and \( \sqrt{24/T} \), which are approximately 0.03 and 0.06 for the observed sample sizes. The high statistical significance of the skewness and kurtosis statistics are in part due to use the daily, as opposed to monthly or quarterly, returns. It is common for the term structure of skewness and kurtosis to decline for many asset returns, which is a strong indicator of outliers.

All of the factors have significant excess kurtosis, even though HML, SMB, and momentum are long-short portfolios. The kurtosis is mitigated by standardized returns, which indicates strong time-variation in volatility (Rosenberg (1972)). All of the portfolios have skewness and kurtosis after standardizing, although it is important
to note that HML has positive skewness, whereas the other portfolios have negative skewness. The positive skewness is puzzling due to the high average returns. While it is common to find assets with high average returns and negative skewness, it is more unusual to find assets with high average returns and positive skewness.

Figure ?? provides an alternative view of the tails of the distribution for each standardized factor return. The top portion of the density is truncated in order to focus on the tails of the distribution. For the market, a number of large outliers are apparent in the left tail. The two largest outliers, the crash in 1987 and the mini-crash in October of 1989, are more than 10 standard deviations from zero. The move in February 2007 was a -7.6 standard deviation move. These moves generate the negative skewness and excess kurtosis.

The upper right panel displays the distribution of normalized returns for SMB. There are both positive and negative outliers, although many more are negative. This can also be seen by the fact that the factor return histogram falls underneath the standard normal density for standardized returns in the +2 to +4 range. The lower right panel displays momentum’s distribution of normalized returns, indicating a large number of negative outliers. The lower left panel displays the results for HML, which has two noticeable features. First, relative to the market and SMB, the range of outcomes for standardized HML returns is less, which generates the relatively lower conditional kurtosis. Second, unlike the other factors, HML has a substantial number of positive outliers, with only a single substantial negative outlier. The positive conditional skewness for HML is an attractive feature (as opposed to a “risk”), and measured by its kurtosis, HML is the least risky of all the portfolios considered. To our knowledge, this feature of HML has not been previously reported in the literature.

2.3.2 Parameter estimates

Table 2.2 and 2.3 summarize the parameter estimates for the four factors. Each table reports posterior means with the posterior standard deviations given directly below
for each of the three main models considered. For ease of discussion, we will refer to “statistical significance” in the traditional way, when the posterior mean is more than two or three standard deviations away from a benchmark value. This is done for ease of communication, instead of reporting full posterior credible sets. Recall, the SV model is a pure stochastic volatility model, SVJ augments the SV model with jumps in returns, and the SVCJ model additionally adds jumps in volatility. The log-likelihood ratio is given at the bottom of each column.

First, consider the model implied behavior of the market. All of these parameter estimates are consistent with previous studies (e.g., Eraker, Johannes, and Polson (2003)) despite the different time spans and data sets (equally weighted market portfolio vs. S&P 500). In all of the models, there is a strong diffusive leverage effect as $\rho$ is strongly negative in all of the models, with magnitudes varying from -48% to -59%. As more shocks are introduced, $\sigma_v$ falls as expected as the normally distributed shocks in volatility play a lesser and lesser role. For the models with jumps, jumps are infrequent (1-2 per year) but relative large with average jump sizes of about -2% and standard deviations about 2% to 3%. In the SVCJ, volatility can move rapidly so jumps in returns tend to be somewhat smaller. An averaged size jump in volatility increases $\sqrt{V_t}$ from about 15% to 23%. Across the models, there is no significant relationship between average returns and volatility, as $\beta$ is sometimes positive, sometimes negative, and not significant in the more general models. The likelihood ratios strongly support the specifications with jumps in returns and in volatility over those without those additional factors.

From this, we obtain a view of aggregate returns as a “carry” trade. In normal or quiet periods, the market delivers high average returns, around 10.5-12%. These high average returns during normal can be viewed as compensation for the jump risk. In periods of market stress, when jumps in returns/volatility occur, returns are quite

---

4 In the SVJ model, the average return implied by the model is $E(r_{t+1}) = \alpha + \beta \left( \frac{\sigma_v}{1-\beta} \right)$, which is about 10.5%.
negative. On average, the jump component contributes an average return of $252 \cdot \lambda \cdot \mu_z$, which is -3.1% in the SVJ model and 3.9%, tempering the high returns in normal times.

SMB returns closely mirror the behavior of the market, with large negative jump sizes, highly persistent volatility, and no systematic relationship between average returns and volatility. Unlike the market, SMB does not have significant "carry," since average returns are slightly negative. The full-sample likelihood ratio statistic indicates that the SVCJ provides a substantially better fit than SVJ, which in turn provides a better fit than the SV model. SMB also has a much smaller leverage effect, as $\rho$ is about -11% to -13% in the SVJ and SVCJ models, respectively.

The value and momentum portfolios are particularly interesting, with the parameter estimates given in Table 2.3. HML has a number of noticeable features. Focusing on the two models with jumps, the first thing to note is that $\rho$ and $\mu_z$ are both positive, although not particularly large in magnitude. This does imply, however, that HML has positive, albeit modest, leverage effects both for normal shocks to volatility and for large shocks in volatility. This implies that as HML's volatility increases, HML delivers above average returns due to the correlations. This is quite different than the results for the market. In the SVJ model, jumps are extremely rare, one every ten years, and are not particularly significant in either an economic or statistical sense. In the SVCJ model, the preferred specification, the mean jump size in returns is 1%. If volatility is at its long run average, 8% annualized, then an averaged size jump in variance would increase volatility to about 16%. At this stage, an interesting question is to understand what causes HML's volatility to increase so rapidly.

Average returns for the value factor are strongly positively related to $V_t$, as measured by $\beta$. In fact, in all there specifications, $\alpha$, which is annualized, is not economically or statistically significant. This indicates that average HML returns vary over time and the variation is related to volatility. This high-frequency variation is distinct from the lower-frequency variation typically measured by valuation ratios or
consumption growth. Together with the shock structure, the value portfolio has a particular striking returns: average returns are positive; when volatility through a normally distributed shock, HML returns are positive on average; when volatility increases through a jump, average returns are positive and economically large (1%, which can be compared with a daily standard deviation of 0.5%); and in the state of the world in which HML's volatility is high, HML pays out higher than normal average returns. Clearly, HML is a great hedge asset and, along with its positive skewness, has all the traits of a statistical arbitrage. Based on the model decompositions, it is hard to identify periods in which HML performs poorly.

The behavior of the momentum factor is different, but similarly interesting. Momentum has very high average returns as measured by \( \alpha \), roughly 15% to 16% per year. Average momentum returns also vary strongly with \( V_t \), but the relationship is negative, indicating that momentum does poorly in periods of high volatility. In terms of leverage effects, momentum returns are modestly positively correlated with the normally distributed shocks in volatility (about 15% to 20%). In the jump models, momentum has significantly negative mean jump sizes, indicate that when volatility increases rapidly, average returns tend to be quite negative.

In contrast to HML, momentum is classic "carry" trade and a poor hedge asset. Momentum delivers high returns in the very low volatility state, and performs poorly when the market is stressed, either via high volatility or a rapid increase in volatility. There is little difference in the likelihood ratios, although the simple SV model appears to fit the best. This will be discussed in greater detail below in Section 3.4.

---

5 This result can also be seen at lower frequencies as noted in Guidolin and Timmerman (2007). They analyze a multivariate regime switching model, and find that there is a high volatility regime in which HML has economically large and statistical significantly positive returns.
2.3.3 Volatility, jump time, and jump size estimates

The output of the MCMC algorithm can also be used to estimate the latent states. There are many practice uses of understanding the dynamics of latent states. For instance, they give valuable information regarding how stochastic volatility and jumps attribute to return variations. From a practitioner's perspective, estimates of volatility enable the asset manager to calculate value-at-risk of a certain portfolio. In this paper, we use the full-information smoothed estimates rather than the real-time filtered estimates, as they provide a more accurate estimate of the state variables at any given point in time.

Figures 2.2 and 2.3 provide time series for each factor's estimated volatility from the SVCJ model. First, let us summarize the volatility of market return as a benchmark. As documented in previous studies, volatility of market return is clearly time-varying, with periods of low and high volatility. Furthermore, there are clear evidence of sudden surges in volatility as represented by episodes in 1987, 1997, 1998 and February of 2007. Now, let us examine the volatility of the asset pricing factors HML, SMB and UMD. Again, the estimated volatility is clearly time-varying, as similar to those of the market returns. Market volatility and size volatility have similar dynamics over time, where market is more volatile and the size factor is smoother. In contrast to the market, the size factor does not identify a jump in volatility in February of 2007. HML and momentum has similar factor volatility dynamics, which are quite different than those of the market and SMB. Another feature of the volatility dynamics are that the factor volatilities seem to identify a jump in 2000 which is the highest in magnitudes compared to the entire time-series, while for the market, it is the third highest episodes in the time-series.

The top panel of Table 2.4 provides correlations between estimated volatilities for the different factors, both in levels and in first differences. Over the full sample, the factor volatilities are significantly correlated in levels, between 57% and 74%. However, the changes in volatilities are less correlated, and in some cases display virtually
no correlation except between market and UMD, and market and HML. This result is very important from an asset management perspective. Trading strategies which are derived from the levels of the second moments of returns provide no diversification benefit in a portfolio. On the other hand, trading strategies which are generated from the changes in factor volatilities have high diversification benefit. This result shows that it is useful to combine signals to form a portfolio.

2.4 Model comparison

Unlike most of the literature, we focus on sequential likelihood ratios, tracking the relative performance of different models over time. In contrast to full-sample tests, this sequential model monitoring provides a unique view of how competing models fit the data on a day-to-day basis. The full sample likelihood estimates were provided in Table 2.2 and 2.3. The full likelihood clearly show that the SVCJ model dominates for market, SMB and HML while it seems that the SV model dominates for momentum.

Figures 2.4 and 2.5 provide time series for each factor's likelihood ratio. Again, as a benchmark, we summarize the results on the market return. As previous research has indicated, the SVCJ model provides the best fit to the data over time. The outperformance stands out especially during periods of market stresses. Whenever our model estimates a jump, the likelihood ratio of SVCJ to SV and SVJ models indicate a surge in the likelihood of the double jump model. The results are mixed for the asset pricing factors. The size factor, SMB, shows very similar likelihood ratio over time as the market. While the value factor, HMB, has a dynamics that is not very similar, the SVCJ model still dominates all other models. Finally, momentum does not show clear dominance of any model. The results are mixed over time.
2.5 Accounting for market returns

The results given in the previous sections did not account for the fact that HML, SMB, and momentum are somewhat correlated with market returns. To account for the market, we consider a simple extension that includes the market as an additional regressor variable:

\[ r_{t+1} = \alpha + \beta V_t + \beta_m r_{t+1}^m + \sqrt{V_t} \varepsilon_{t+1} + J_{t+1} Z_{t+1}, \]

where \( r_{t+1}^m \) is the market factor. This purges the factor's of the contemporaneous correlation with the market. This specification has a CAPM-like intuition.

Table 2.5 reports parameter estimates for each of the model specifications for the size, value, and momentum portfolios. While most of the estimates have similar intuition as the parameter estimates from before, there are some key differences. After accounting for the market, HML expected return \( \alpha \) becomes significant and jump mean becomes negative, but insignificant. Due to the insignificance of the negative jump mean, HML remains like a "statistical arbitrage" from a practitioner's perspective. After adjusting for market return, SMB also has a similar effect that expected return \( \alpha \) becomes positive and significant. In contrast to HML and SMB, momentum's estimates remain relatively the same. The "carry" like feature of momentum remains. Finally, in all cases, jumps are more rare after accounting for the market, indicating that some of the periods of market stresses were systematic risk and could not be hedged away.

The volatility estimates are very similar to those in Figures 2.4 and 2.5 and are thus not reported here. We summarize a few key differences here. For HML and momentum, the jump mean of volatility decreased while for SMB, the jump mean of volatility increased. For all three factors, the jump mean of volatility remains significant. This is a very interesting results. HML and momentum have an idiosyncratic risk component which is lower once you use the market portfolio to hedge away
some of the underlying risk. On the other hand, the market portfolio did not add diversification benefit to the SMB portfolio.

The bottom panel of Table 2.4 provides correlations between estimated volatilities for the different factors, both in levels and in first differences. Over the full sample, the factor volatilities remain significantly correlated in levels, between 60% and 75%. While the correlation of first difference of factor volatilities between market return and the SMB portfolio remains low, the changes in volatilities become much more correlated after accounting for market influences.

As a metric for model comparison, we use the full likelihood statistics reported in Table 2.2, 2.3 and 2.5. After accounting for market influences, the preferred model for SMB remains to be SVCJ, but the preferred model for HML and momentum is the SV model. Since most of the large movements in the asset pricing factors occurred on the same days as the market, most of the large movements were market influences. After accounting for the market, estimated jump intensity becomes lower and jump mean in some cases become insignificant. Hence, the preferred model becomes the SV model. To fully test this hypothesis, an out-of-sample forecasting exercise will be useful to identify the best performing model. However, this is not pursued in this paper. Comparing the lowest likelihood from the market model with their counterparts from the earlier section, we find that the market model is preferred in all cases. Viewing from a CAPM perspective, this should not be a surprise as the market portfolio provides diversification benefits.

### 2.6 Conclusions

This paper provides an alternative view of common asset pricing factors. By analyzing the time-series properties of factor returns and decomposing their variations into various components, we provide a new set of stylized facts about common asset pricing factors. In particular, two results are particularly striking. First, the fact that the volatility of HML and UMD are so variable over time and at high frequencies.
Second, the very strong relationships between HML and UMD average returns and their volatility. These results pose challenges to theories that attempt to explain the returns on these portfolios. In addition to explaining the high average returns and low volatility, these theories must also address the strong, high frequency variation in the risks in the portfolios, and the state-dependent payoff patterns of these portfolios.

From an empirical perspective, this paper is the first to consider the daily performances of various models to fit the asset pricing factors. Finally, as a future research agenda, an out-of-sample forecasting exercise will another valuable metric to assess the performance of the various models analyzed in this paper.

2.7 Appendix: Algorithmic details

2.7.1 MCMC algorithms

This section provides details regarding the MCMC algorithm. The algorithm cycles through the posterior conditionals of the parameters and state variables. If we denote \( \theta = (\theta_1, ..., \theta_k) \) as a partition of the parameters, \( \theta_{-i} = (\theta_1, ..., \theta_{i-1}, \theta_{i+1}, ..., \theta_k) \) as the vector of parameters not including \( \theta_i \), and the latent states as \( X = (J, V, Z) \) where, e.g., \( J = (J_1, ..., J_T) \). Generically, the algorithm is given by cycling through the following conditional distributions:

\[
p(\theta_1|x, x, r)
\]
\[
p(\theta_2|x, x, r)
\]
\[
\vdots
\]
\[
p(\theta_k|x, x, r)
\]
\[
p(J|x, V, Z, r)
\]
\[
p(Z|x, V, J, r)
\]
\[
p(V_t|x, V_{-t}, Z, J, r) \text{ for } t=1, ..., T.
\]
The parameters will be blocked into groups, as shown below. Notice for updating volatility, we use single state updating and, as discussed below, the jump times and sizes are conditionally independent over time so they can be updated one at a time.

We now discuss the blocking and sampling from the complete conditionals. For the parameters \((\alpha, \beta)\), we write the returns equation as

\[
\tilde{r}_{t+1} = \frac{r_{t+1} - J_{t+1}Z_{t+1}}{\sqrt{V_t}} - \sigma_v \sqrt{1 - \rho^2} \tilde{e}_{t+1}^v = \beta_r X_t + \sigma_v \rho \tilde{e}_{t+1},
\]

where \(X_t = \left[ \frac{1}{\sqrt{V_t}}, \sqrt{V_t} \right] \) is a vector of regressors,

\[
\tilde{e}_{t+1}^v = \frac{V_{t+1} - \alpha_v - \beta_v V_t - J_{t+1}Z_{t+1}^v}{\sigma_v},
\]

and \(\tilde{e}_{t+1} \sim \mathcal{N}(0,1)\) is i.i.d. standard normal. The two error terms arise by writing normal error in the returns equation as

\[
\tilde{e}_{t+1}^v = \rho \tilde{e}_{t+1}^v + \sqrt{(1 - \rho^2)} \tilde{e}_{t+1}^v
\]

and then solving for \(\tilde{e}_{t+1}^v\) in the variance equation. Assuming a normal prior distribution, \(\beta_r \sim \mathcal{N}(a, A)\), then

\[
p(\beta_r | \tilde{r}^T, V^T, \Theta, A, a) \sim \mathcal{N}(a_T, A_T)
\]

where \(a_T\) and \(A_T\) are straightforward to compute using standard conjugate updating. In practice we chose \(a = 0\) and \(A = 1\), generating an uninformative prior as the the prior is quite flat over the relevant parameter range.

Next, for the parameters \(\alpha_v, \beta_v, \sigma_v\), and \(\rho\), we reparameterize the model, as in Jacquier, Polson, and Rossi (2004). Defining \(\Omega = \sigma_v \sqrt{1 - \rho^2}\) and \(\psi = \sigma_v \rho\), it is clear that \((\alpha_v, \beta_v, \sigma_v, \rho)\) is equivalent to \((\alpha_v, \beta_v, \psi, \Omega)\). The motivation for the reparameter-
terization is that the original volatility evolution

\[ V_{t+1} = \alpha_v + \beta_v V_t + \rho \sigma_v \sqrt{V_t} \epsilon_{t+1} + \sigma_v \sqrt{(1 - \rho^2) \sqrt{V_t} \epsilon_{t+1}^v}, \]

where \( \epsilon_{t+1}^v \) is i.i.d. normal, can be re-written as

\[ \frac{V_{t+1}}{\sqrt{V_t}} = \frac{\alpha_v}{\sqrt{V_t}} + \beta_v \sqrt{V_t} \psi \epsilon_{t+1} + \Omega \epsilon_{t+1}^v \]

\[ V_{t+1}^* = \beta_v^* X_t^v + \Omega \epsilon_{t+1}^v, \]

where

\[ X_t^v = \left( \frac{1}{\sqrt{V_t}}, \sqrt{V_t}, \epsilon_{t+1} \right) \]

and

\[ \epsilon_{t+1} = \frac{r_{t+1} - J_{t+1} Z_{t+1} - \alpha - \beta V_t}{\sigma_v \sqrt{V_t}} \]

is observed. This is just a standard multivariate regression model. For the priors, we assume conjugate priors For the priors, we assume \( (\alpha_v, \beta_v, \psi | \Omega) \sim N(b, \Omega B), p(\Omega) \sim IG(c, C) \), which leads to conjugate posteriors:

\[ p(\Omega^2 | V^*, X^v) \sim IG(\sigma_T, C_T) \]

\[ p(\beta_v^* | \Omega, V^*, X^v) \sim N(b_T, \Omega^2 B_T), \]

where the hyperparameters are easy to compute using standard regression updating.

Updating \((\mu_v, \sigma_v^2)\) and \(\lambda\) are also straightforward. We assume \(\mu_v \sim N(d, D)\) and \(\sigma_v^2 \sim IG(e, E)\). The prior on \(\mu_v\) is uninformative, with \(d = 0\) and \(D = 1\), and update the parameters by marginalizing out the jump sizes. To see this, write

\[ r_{t+1} = \alpha + \beta V_t + \sigma_v \sqrt{V_t} + J_{t+1} (\mu_v + \sigma_v \varepsilon_{t+1}^I) \]

where \(\varepsilon_{t+1}^I\) is i.i.d. normal. The conditional posterior for \(\mu_v\) is normal, but the
posterior for $\sigma_z$ is not a known distribution. We use an independence Metropolis-Hasting updating step with an inverse Gamma proposal that bounds the tails of the target density. The only informative prior that we use is on $\sigma_z^2$, where we specify that $E(\sigma_z^2) = (3\text{std}(r))^2$ and $\text{var}(\sigma_z^2) = (3\text{std}(r))^2$. This bounds $\sigma_z$ away from zero, avoiding the identification issues that are present in the mixture models.
<table>
<thead>
<tr>
<th>Factor</th>
<th>Market</th>
<th>SMB</th>
<th>HML</th>
<th>UMD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average return (annualized, %)</td>
<td>7.48</td>
<td>-0.18</td>
<td>4.63</td>
<td>10.16</td>
</tr>
<tr>
<td>Standard deviation (daily,%)</td>
<td>0.96</td>
<td>0.55</td>
<td>0.50</td>
<td>0.65</td>
</tr>
<tr>
<td>Standard deviation (annualized,%)</td>
<td>15.27</td>
<td>8.78</td>
<td>7.96</td>
<td>10.30</td>
</tr>
<tr>
<td>Average return/standard deviation</td>
<td>0.49</td>
<td>-0.02</td>
<td>0.58</td>
<td>0.99</td>
</tr>
<tr>
<td>Skewness (raw returns)</td>
<td>-0.96</td>
<td>-1.44</td>
<td>0.05</td>
<td>-0.98</td>
</tr>
<tr>
<td>Kurtosis (raw returns)</td>
<td>19.19</td>
<td>30.70</td>
<td>6.76</td>
<td>11.41</td>
</tr>
<tr>
<td>Skewness (standardized returns)</td>
<td>-0.63</td>
<td>-0.27</td>
<td>0.22</td>
<td>-0.42</td>
</tr>
<tr>
<td>Kurtosis (standardized returns)</td>
<td>5.24</td>
<td>1.89</td>
<td>1.51</td>
<td>2.21</td>
</tr>
</tbody>
</table>

Table 2.1: Summary statistics for asset pricing factors. For each factor, summary statistics are computed using daily data from January 1980 through December 2007, obtained from Ken French’s website. The kurtosis statistics are excess kurtosis. Standardized returns are computed as raw returns divided by the previous 22-day equally weighted volatility.
Table 2.2: Parameter estimates for the market and SMB. The parameters $\alpha$ and $\lambda$ are annualized, all other parameters are in daily units.

<table>
<thead>
<tr>
<th></th>
<th>Market SV</th>
<th>Market SVJ</th>
<th>Market SVCJ</th>
<th>SMB SV</th>
<th>SMB SVJ</th>
<th>SMB SVCJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>12.5058</td>
<td>8.1436</td>
<td>11.9551</td>
<td>-1.3152</td>
<td>-0.2225</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.7565)</td>
<td>(2.2332)</td>
<td>(1.6503)</td>
<td>(2.0563)</td>
<td>(2.0019)</td>
<td>(1.9623)</td>
</tr>
<tr>
<td>$\beta$</td>
<td>-0.0234</td>
<td>0.0109</td>
<td>-0.0035</td>
<td>-0.0529</td>
<td>0.0261</td>
<td>0.0116</td>
</tr>
<tr>
<td></td>
<td>(0.0036)</td>
<td>(0.0119)</td>
<td>(0.0096)</td>
<td>(0.0288)</td>
<td>(0.0302)</td>
<td>(0.0288)</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>—</td>
<td>1.2255</td>
<td>2.2877</td>
<td>—</td>
<td>0.5373</td>
<td>0.6101</td>
</tr>
<tr>
<td></td>
<td>(0.4168)</td>
<td>(0.6912)</td>
<td>(0.2200)</td>
<td>(0.1938)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_z$</td>
<td>—</td>
<td>-2.5270</td>
<td>-1.7242</td>
<td>—</td>
<td>-1.3120</td>
<td>-1.3232</td>
</tr>
<tr>
<td></td>
<td>(0.8659)</td>
<td>(0.4359)</td>
<td>(1.1356)</td>
<td>(0.9619)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_z$</td>
<td>—</td>
<td>2.9576</td>
<td>1.9487</td>
<td>—</td>
<td>3.5082</td>
<td>3.3315</td>
</tr>
<tr>
<td></td>
<td>(0.6249)</td>
<td>(0.3282)</td>
<td>(0.7692)</td>
<td>(0.6515)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho$</td>
<td>-0.4832</td>
<td>-0.5688</td>
<td>-0.5860</td>
<td>0.0024</td>
<td>-0.1111</td>
<td>-0.1332</td>
</tr>
<tr>
<td></td>
<td>(0.0371)</td>
<td>(0.0390)</td>
<td>(0.0411)</td>
<td>(0.0519)</td>
<td>(0.0567)</td>
<td>(0.0644)</td>
</tr>
<tr>
<td>$\mu_v$</td>
<td>—</td>
<td>—</td>
<td>1.1261</td>
<td>—</td>
<td>—</td>
<td>0.6086</td>
</tr>
<tr>
<td></td>
<td>(0.2033)</td>
<td>(0.2050)</td>
<td>(0.2050)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_v$</td>
<td>0.0222</td>
<td>0.0147</td>
<td>0.0158</td>
<td>0.0080</td>
<td>0.0048</td>
<td>0.0062</td>
</tr>
<tr>
<td></td>
<td>(0.0009)</td>
<td>(0.0017)</td>
<td>(0.0011)</td>
<td>(0.0010)</td>
<td>(0.0007)</td>
<td>(0.0010)</td>
</tr>
<tr>
<td>$\beta_v$</td>
<td>0.9738</td>
<td>0.9814</td>
<td>0.9682</td>
<td>0.9716</td>
<td>0.9820</td>
<td>0.9721</td>
</tr>
<tr>
<td></td>
<td>(0.0011)</td>
<td>(0.0021)</td>
<td>(0.0028)</td>
<td>(0.0033)</td>
<td>(0.0025)</td>
<td>(0.0046)</td>
</tr>
<tr>
<td>$\sigma_v$</td>
<td>0.1503</td>
<td>0.1168</td>
<td>0.1033</td>
<td>0.0768</td>
<td>0.0550</td>
<td>0.0521</td>
</tr>
<tr>
<td></td>
<td>(0.0060)</td>
<td>(0.0074)</td>
<td>(0.0078)</td>
<td>(0.0053)</td>
<td>(0.0045)</td>
<td>(0.0050)</td>
</tr>
<tr>
<td>Log L</td>
<td>-8654.17</td>
<td>-8640.52</td>
<td>-8611.64</td>
<td>-5019.51</td>
<td>-4972.73</td>
<td>-4954.72</td>
</tr>
<tr>
<td></td>
<td>HML</td>
<td></td>
<td>UMD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-----------</td>
<td>-------</td>
<td>-----------</td>
<td>-------</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SV</td>
<td>SVJ</td>
<td>SVCJ</td>
<td>SV</td>
<td>SVJ</td>
<td>SVCJ</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>-0.0549</td>
<td>-0.0214</td>
<td>0.0504</td>
<td>15.4519</td>
<td>15.1029</td>
<td>16.6544</td>
</tr>
<tr>
<td></td>
<td>(1.1688)</td>
<td>(1.0700)</td>
<td>(0.8394)</td>
<td>(0.9281)</td>
<td>(0.1772)</td>
<td>(1.3613)</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.0774</td>
<td>0.0761</td>
<td>0.0598</td>
<td>-0.0546</td>
<td>-0.0319</td>
<td>-0.0533</td>
</tr>
<tr>
<td></td>
<td>(0.0179)</td>
<td>(0.0136)</td>
<td>(0.0096)</td>
<td>(0.0051)</td>
<td>(0.0146)</td>
<td></td>
</tr>
<tr>
<td>$\lambda$</td>
<td>—</td>
<td>0.1068</td>
<td>0.8697</td>
<td>—</td>
<td>1.3283</td>
<td>1.0226</td>
</tr>
<tr>
<td></td>
<td>(0.1903)</td>
<td>(0.2389)</td>
<td>(0.5249)</td>
<td>(0.2991)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_z$</td>
<td>—</td>
<td>0.0997</td>
<td>1.0111</td>
<td>—</td>
<td>-1.6870</td>
<td>-1.5425</td>
</tr>
<tr>
<td></td>
<td>(8.0480)</td>
<td>(0.3566)</td>
<td>(0.5442)</td>
<td>(0.3920)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_z$</td>
<td>—</td>
<td>1.3789</td>
<td>0.9348</td>
<td>—</td>
<td>1.5082</td>
<td>1.2833</td>
</tr>
<tr>
<td></td>
<td>(0.4872)</td>
<td>(0.1692)</td>
<td>(0.3131)</td>
<td>(0.2334)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.0922</td>
<td>0.0823</td>
<td>0.0330</td>
<td>0.1576</td>
<td>0.1657</td>
<td>0.1913</td>
</tr>
<tr>
<td></td>
<td>(0.0387)</td>
<td>(0.0392)</td>
<td>(0.0436)</td>
<td>(0.0372)</td>
<td>(0.0358)</td>
<td>(0.0407)</td>
</tr>
<tr>
<td>$\mu_u$</td>
<td>—</td>
<td>—</td>
<td>0.8316</td>
<td>—</td>
<td>—</td>
<td>1.1204</td>
</tr>
<tr>
<td></td>
<td>(0.2199)</td>
<td>(0.2871)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_u$</td>
<td>0.0028</td>
<td>0.0029</td>
<td>0.0031</td>
<td>0.0043</td>
<td>0.0032</td>
<td>0.0049</td>
</tr>
<tr>
<td></td>
<td>(0.0004)</td>
<td>(0.0004)</td>
<td>(0.0003)</td>
<td>(0.0003)</td>
<td>(0.0000)</td>
<td>(0.0005)</td>
</tr>
<tr>
<td>$\beta_u$</td>
<td>0.9881</td>
<td>0.9880</td>
<td>0.9766</td>
<td>0.9891</td>
<td>0.9915</td>
<td>0.9753</td>
</tr>
<tr>
<td></td>
<td>(0.0016)</td>
<td>(0.0017)</td>
<td>(0.0022)</td>
<td>(0.0009)</td>
<td>(0.0002)</td>
<td>(0.0023)</td>
</tr>
<tr>
<td>$\sigma_u$</td>
<td>0.0587</td>
<td>0.0590</td>
<td>0.0495</td>
<td>0.0799</td>
<td>0.0684</td>
<td>0.0705</td>
</tr>
<tr>
<td></td>
<td>(0.0042)</td>
<td>(0.0042)</td>
<td>(0.0036)</td>
<td>(0.0035)</td>
<td>(0.0030)</td>
<td>(0.0051)</td>
</tr>
<tr>
<td>Log L</td>
<td>-3575.93</td>
<td>-3585.99</td>
<td>-3571.54</td>
<td>-4842.59</td>
<td>-4855.70</td>
<td>-4850.73</td>
</tr>
</tbody>
</table>

Table 2.3: Parameter estimates for the HML and momentum. The parameters $\alpha$ and $\lambda$ are annualized, all other parameters are in daily units.

<table>
<thead>
<tr>
<th></th>
<th>Market</th>
<th>SMB</th>
<th>HML</th>
<th></th>
<th>Market</th>
<th>SMB</th>
<th>HML</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levels</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMB</td>
<td>74.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HML</td>
<td>65.1</td>
<td>57.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UMD</td>
<td>72.4</td>
<td>61.0</td>
<td>71.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Changes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMB</td>
<td>65.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HML</td>
<td>60.5</td>
<td>66.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UMD</td>
<td>71.9</td>
<td>64.4</td>
<td>75.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2.4: Volatility correlations, in levels and in first differences for the SVCJ model for each factor.
<table>
<thead>
<tr>
<th></th>
<th>SMB</th>
<th></th>
<th>HML</th>
<th></th>
<th>Momentum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SV</td>
<td>SVJ</td>
<td>SV</td>
<td>SVJ</td>
<td>SV</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>5.0357</td>
<td>2.6861</td>
<td>3.7633</td>
<td>5.8666</td>
<td>4.4962</td>
</tr>
<tr>
<td></td>
<td>1.5201</td>
<td>1.9359</td>
<td>1.6529</td>
<td>1.2124</td>
<td>1.2083</td>
</tr>
<tr>
<td>$\beta$</td>
<td>-0.0489</td>
<td>-0.0024</td>
<td>-0.0148</td>
<td>0.0379</td>
<td>0.0579</td>
</tr>
<tr>
<td></td>
<td>0.0225</td>
<td>0.0298</td>
<td>0.0246</td>
<td>0.0303</td>
<td>0.0333</td>
</tr>
<tr>
<td>$\beta_m$</td>
<td>-0.2562</td>
<td>-0.2521</td>
<td>-0.2543</td>
<td>-0.2609</td>
<td>-0.2613</td>
</tr>
<tr>
<td></td>
<td>0.0045</td>
<td>0.0051</td>
<td>0.0042</td>
<td>0.0032</td>
<td>0.0035</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.3006</td>
<td>0.3578</td>
<td>—</td>
<td>0.4539</td>
<td>0.5358</td>
</tr>
<tr>
<td></td>
<td>0.1772</td>
<td>0.1383</td>
<td>0.2523</td>
<td>0.1882</td>
<td>—</td>
</tr>
<tr>
<td>$\mu_z$</td>
<td>-3.7252</td>
<td>-3.2863</td>
<td>—</td>
<td>-0.5478</td>
<td>-0.7876</td>
</tr>
<tr>
<td></td>
<td>1.8298</td>
<td>1.1419</td>
<td>0.9371</td>
<td>0.5273</td>
<td>—</td>
</tr>
<tr>
<td>$\sigma_z$</td>
<td>2.7501</td>
<td>2.813</td>
<td>—</td>
<td>1.6809</td>
<td>1.5167</td>
</tr>
<tr>
<td></td>
<td>0.7371</td>
<td>0.6194</td>
<td>0.4351</td>
<td>0.3276</td>
<td>—</td>
</tr>
<tr>
<td>$\rho$</td>
<td>-0.0972</td>
<td>-0.1435</td>
<td>-0.1424</td>
<td>0.0113</td>
<td>0.0164</td>
</tr>
<tr>
<td></td>
<td>0.0369</td>
<td>0.0420</td>
<td>0.0432</td>
<td>0.0371</td>
<td>0.0384</td>
</tr>
<tr>
<td>$\mu_v$</td>
<td>1.1137</td>
<td>—</td>
<td>—</td>
<td>0.5314</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>0.4074</td>
<td>0.1688</td>
<td>0.1688</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>$\alpha_v$</td>
<td>0.0063</td>
<td>0.0045</td>
<td>0.0055</td>
<td>0.0031</td>
<td>0.0028</td>
</tr>
<tr>
<td></td>
<td>0.0006</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0002</td>
<td>0.0003</td>
</tr>
<tr>
<td>$\beta_v$</td>
<td>0.9762</td>
<td>0.9825</td>
<td>0.9739</td>
<td>0.9803</td>
<td>0.9820</td>
</tr>
<tr>
<td></td>
<td>0.0022</td>
<td>0.0021</td>
<td>0.0021</td>
<td>0.0015</td>
<td>0.0019</td>
</tr>
<tr>
<td>$\sigma_v$</td>
<td>0.0810</td>
<td>0.0655</td>
<td>0.0658</td>
<td>0.0569</td>
<td>0.0528</td>
</tr>
<tr>
<td></td>
<td>0.0039</td>
<td>0.0040</td>
<td>0.0040</td>
<td>0.0029</td>
<td>0.0030</td>
</tr>
<tr>
<td>$L$</td>
<td>-4493</td>
<td>-4479</td>
<td>-4458</td>
<td>-2476</td>
<td>-2480</td>
</tr>
</tbody>
</table>

Table 2.5: Parameter estimates for SMB, HML and momentum taking accounting for market returns. The parameters $\alpha$ and $\lambda$ are annualized, all other parameters are in daily units.
Figure 2.1: Histograms for standardized factor returns. The solid red bars provide a histogram of the factor returns that are standardized by previous one-month's volatility and normalized to have a mean of zero and standard deviation of one. The solid black line is the density of a standard normal random variable.
Figure 2.2: Time series of estimated volatilities. For each factor, the figure displays the posterior mean of volatility.
Figure 2.3: Time series of estimated volatilities. For each factor, the figure displays the posterior mean of volatility.
Figure 2.4: Time series of likelihood ratio.
Figure 2.5: Time series of likelihood ratio.
Chapter 3

Lending Without Access to Collateral: A Theory of Micro-Loan Borrowing Rates & Defaults

3.1 Introduction

There are very large groups of society, especially in poor and developing parts of the world who do not have access to rudimentary financial services such as bank savings accounts, credit facilities, or insurances. Households in these sections of the society are typically poor and access credit in informal credit markets. Such informal credit markets include: a) local money-lenders, b) local shop-keepers, who provide trade credit, c) pawn-brokers, d) payday lenders, and e) Rotating Savings and Credit Associations (ROSCAS). A number of economists have examined these informal credit markets, and their potential linkages to more formal credit markets. A partial list of such research includes Besley, Coate, and Loury (1993), Braverman, and Guasch (1986), Varghese (2000, 2002), and Caskey (2005). It is well understood
that the interest rates in such informal markets tend to be much higher than the borrowing rates that prevail in formal credit markets$^1$.

Micro-loan markets represent one of the more recent developments, which enable poor households to access credit. These are markets where very small (hence micro) loans are extended to poor households. Often, such loans are given only to women, and in groups. Borrowers in these markets have no meaningful physical collateral and are heavily credit constrained. Micro-loans are characterized by three essential features: a) loans are short-term in nature, relatively small amounts and consummated without physical collateral, but structured with social collateral$^2$; b) loans are extended typically to a group, whose size can range from five (in the Grameen model) to twenty (in the Self-Help-Groups or SHG), where the group members are jointly liable for default by any member of the group; and c) loans carry frequent interest payments (weekly in many cases) and carry significant administrative expenses that are incurred in order to ensure timely delivery of loans to remote villages and for the collection of payments$^3$.

To our understanding, no formal model has been developed for understanding the determination of borrowing rates and default rates in micro loans, and how they depend on various features of the micro-loan contracts$^4$; nor has there been a treatment of the relationship between the market structures of loans (competitive or monopolistic) and the equilibrium borrowing rates and default rates. This is the primary objective of the paper. The paper is of a broader interest since it delivers a framework to think about enforcing loan contracts when the lender has no access to physical

---

$^1$ A complete survey of research in informal credit markets and micro-finance is well beyond the scope of our paper. We refer to two excellent sources: 1) Armendariz, and Morduch (2005), and 2) Bolton and Rosenthal (2005).

$^2$ Defaulting borrowers often have to live in the same community in front of whom they defaulted; this leads to a sense of shame leading sometimes to extreme outcomes.

$^3$ See Shankar (2006) for an analysis of transactions costs in micro credit.

$^4$ See The Economics of Microfinance (Armendariz and Morduch), Forthcoming from the MIT Press, 2005, for a full discussion of this area.
Since 1976, micro-finance and micro-loans have emerged as a sector where poor households are able to accumulate savings and access credit. Additional financial services such as rainfall insurance, livestock insurance, and health insurance are also being provided increasingly through these channels. Table 3.1 illustrates the size of the micro-loan market as of 2003, based on voluntary reporting of lending institutions to a centralized database maintained by MIX.

The total size of the micro-loan market covering a little over 28 million borrowers as reported in Table 3.1 is about $8.7 billions; this is potentially a very serious underestimate of the actual size of the market since many lenders do not report their activities to MIX. Another estimate found in Microcredit Summit (2003) reports that nearly 2500 lending institutions covered a total of 67 million borrowers as of 2002. Note that Latin America and East Asia are the regions that account for more than 50% of the loans, but South Asia accounts for more than 50% of active borrowers. The total number of active borrowers based on voluntarily reported data is in excess of 28 millions. Regardless of the estimates, what is clear is that the number of households in need of rudimentary financial services such as loans, savings, and insurance is considerably higher. For example, in India alone, the estimated number of people in need of rudimentary financial services is over 200 millions.

Morduch (1999) provides estimates of borrowing rates in this market, but no systematic evidence of defaults is available to our knowledge. In this section, we use the MIX data to estimate interest rates on micro-loans, ex-ante assessment of default exposure, and ex-post write-offs on loans.

Table 3.1 also provides estimates of borrowing rates by assuming that the net revenue reported by the institutions are comprised exclusively by loan interest income.

---

5 See Ananth, Barooah, Ruchismita, and Bhatnagar (2004) for an illuminating discussion on designing a framework for delivering financial services to the poor in India.

6 MIX was incorporated in June 2002 as a not-for-profit private organization. The MIX (Microfinance Information eXchange) aims to promote information exchange in the micro-finance industry.
only. This is likely to be a somewhat noisy estimate because the income may include both principal and interest payments, as well as income from other sources such as investments made by the lending institutions. Hence the estimates of interest rates reported in Table 3.1 should be interpreted with caution. The ex-ante assessment of default is captured by the Portfolio At Risk (PAR) measure reported by each lending institution. The ex-post measure is captured by the write-offs reported in the data set.

The average figures indicate that the micro-loan rates are in excess of 30%. The rates are significantly lower in South Asia, which is characterized by many borrowers who borrow very small amounts. Morduch (1999) has provided estimates of borrowing rates ranging from 20% in Grameen Bank in India to 55% in Indonesia. Our estimates, which reflect the bias of self-reporting institutions, range from 20% to 40%. The rates are much higher in Africa. The rates charged in micro-loans appear to be well below the rates that local money lenders tend to charge. If we think of local money lenders as the outside borrowing option available to the borrowers, then micro-loan rates may not look usurious. The portfolio at risk (PAR) is highest in South Asia, but the write-offs are higher in Africa and Latin America.

It is interesting to note that the PAR estimates are systematically higher than the actual write-offs. This suggests that, ex-ante, the lender believes that the probability of default is much higher than ex-post observed default probability.

To get an appreciation of the composition of lenders, Table 3.2 provides a breakdown of the micro-loans across different lenders. Of the 613 lending institutions voluntarily reporting, banks account for more than 50% of the dollar value of loans, responsible for more than 31% of all active borrowers. In sharp contrast, NGOs account for just 17.3% of the dollar value of the loans, but spans more than 46% of all active borrowers. NGOs and non-bank financial institutions have a similar overall coverage globally with very important cross-sectional variations.

The small size of the loan and the presence of numerous borrowers make in-
vestment, in additional screening and monitoring efforts, an expensive proposition. This makes micro-loan portfolio unattractive for many big commercial banks. Due to these factors, micro-lending approaches focus on a) contractual arrangements, b) punishment conditional on default, c) peer group efforts to in effect substitute social collateral for physical collateral, and d) partnering with local financial institutions, which may possess informational advantages and thus may be able to monitor the loans better.

In Table 3.2, we also provide a breakdown of interest rates and default measures across the different organizational forms. Banks appear to charge the lowest interest rates. The rates charged by other lenders vary, ranging from 30% to 38%.

To summarize, the following facts emerge from our analysis: first, the micro-loan interest rates are rather high, ranging from 30% to nearly 40%. Second, the actual losses as conveyed by the write-off ratios are relatively small, ranging from 0.50% to 3.50%. The lenders tend to estimate the portfolio at risk at a much higher level, averaging around 7%. There is a cross-sectional variation in the borrower size and default rates depending on the organizational arrangement of the lender.

In our numerical illustrations, we will attempt to calibrate the model so that loan rates roughly match these numbers. Then, we will ask how various terms of the loan contract can be used to bring the loan rates down without significantly increasing the default probabilities. In particular, we find that decreasing monitoring has the most significant impact.

The borrowing rates in micro-loan contracts must depend on the following important factors:

1. *Administrative and monitoring expenses:* these are needed to deliver the loans at the doorsteps of poor borrowers in rural areas. High administrative expenses (arguably) keep the default rates low, but render the borrowing costs very high. Note that these costs are borne by the lender, and are eventually passed on to
the micro-borrower. CGAP(2003)\textsuperscript{7} reports that the administrative expenses range from 18.9\% in Asia to 38.2\% in Africa on the loans. Consequently, micro-loan interest rates are rather high; this should, however, be put in the context of the fact that the other "outside options" for the micro-borrowers are even more expensive. For example, CGAP (2003) reports money-lenders charging anywhere from 10\% per month. In Philippines, the estimated daily interest rates on loans made by local money lenders is 20\% per day.

2. \textit{Joint-Liability Arrangements}: when properly structured, groups of borrowers are formed through an assortative matching procedure to provide sufficient peer pressure and monitoring in order to keep the default rates of group members low\textsuperscript{8}. This should reduce the costs of borrowing. Note that the peer monitoring costs are borne by the borrowers. Thus, there is a trade-off for the group: active peer monitoring reduces defaults and delinquencies but increases the efforts required. As noted by Stiglitz (1990) peer monitoring can help to mitigate, if not solve, the ex-ante moral hazard problem: it prevents any member of the group from taking risky projects because others in the group, who are jointly liable will attempt to prevent that from happening. Considerable research has focussed on joint-liability contracting (See Ghatak (1999), and Ghatak and Guinnane (1999), for example.), where borrowers form groups and members of the groups agree to take responsibility for delinquencies and defaults by individual group members. Requiring each member of the group to be liable for the entire group’s liabilities makes peer monitoring much more effective. In addition, the group’s ability to obtain additional loans will be predicated on the entire group fulfilling the contractual obligations on existing loans.

3. \textit{Credible Threat of Punishment Upon default}: the lender must be able to com-

\textsuperscript{7} CGAP stands for Consultative Group to Assist the Poor.

\textsuperscript{8} In assortatively matched groups, potential borrowers form their own group without any intervention from the lender.
municate credibly, ex-ante, that default by the group will lead to significant costs to the group, including the inability to access further loans. This can be done in two ways: first, the target group is chosen so that they have little or no access to formal credit markets, and second, institutional mechanisms such as credit bureaus be put in place so that lenders are able to share information about defaulters and collectively enforce punishment. Furthermore, the group-lending mechanism places a very high social cost on individual defaulters. Often, borrowers are promised additional loans if they successfully pay off existing loans. Rational borrowers know that their access to micro-loans in the future is conditional on not defaulting on existing loans. This provides strong incentives for the micro-loan borrowers not to default.

4. Informational Asymmetry: One aspect of the micro-loans is that banks with depth in lending ability do not necessarily possess informational advantages about many borrowers located in different corners of the country. They also lack the monitoring technology that is needed to enforce payment should joint-liability contracting fail to deliver acceptable recovery rates on loans. The informational advantages and monitoring tools in such a market are usually the domain of local lenders and Micro finance institutions (MFIs). Increasingly, micro-loans are structured with complementary participation by local lenders and bigger financial institutions (see Nini (2004), in the context of emerging market lending): big banks will lend requisite amount of capital to local MFIs at a certain rate of interest. Local MFIs will then re-lend the capital to groups of local borrowers under a joint-liability scheme. The local MFI will assume the “first loss tranche” (say, the initial 10% of losses experienced by the loan portfolio), and the rest will pass through to the bank. In models used by certain banks in India, the loans remain on the books of the bank and not with the local MFI. This so-called partnership model is discussed in Ananth (2005) and Harper and Kirsten (2006).
The primary goal of the paper is to construct a model of borrowing and lending where the lender has no access to collateral and the borrower is severely credit constrained. We first wish to study the impact of market structure on borrowing rates and defaults. In addition, our model relates equilibrium borrowing rates and default probabilities to some salient features of the loan contracts we described earlier. While the model is applied to micro-loans, it is easy to explore other markets in which borrowers obtain loans from lenders who do not have access to collateral upon default. We wish to understand the relative importance of a) lender monitoring, b) punishment upon default, c) maturity structure of the loan, and d) peer monitoring on the loan rates. In particular, we explore whether a particular combination of these variables can keep the default risk low, and at the same time provide lower interest rates to the borrower. Such an outcome should be of great interest in the design of micro-loan contracts. We also wish to conduct a welfare analysis. To this end, we wish to characterize the borrower's value function as a function of these levers that the lender has at her disposal and as a function of the market structure.

In order to keep the problem tractable we have abstracted from the important questions of adverse selection and moral hazard problems and focus directly after the borrowers have formed a group. As a first approximation, we assume that lenders and borrowers have the same information about the cash flows generated by the technology of the borrowing group.

The paper is organized as follows: Section 2 formulates the basic model. In section 3, we characterize the equilibrium and its properties. In particular, we show monitoring does not work in equilibrium, and that the debt maturity can be used to control default probability. We also characterize a) default probabilities b) loan rates, and c) borrower's welfare for varying levels of monitoring, joint-liability efforts, and punishment technology. Section 5 concludes. We also include a technical appendix including the details of the derivations for interested readers.
3.2 Model Specification

Our model is directly specified at the level of the borrowing group. In doing so, we abstract from several interesting questions about how the group is formed, and the role that joint liability plays in the choice of the members of the group as well as the choice of the riskiness of projects by the members of the group. See Stiglitz (1991) and Ghatak (1999) for a treatment of these issues. The salient features of the model are that a) the group cannot undertake any productive investment in the absence of the loan (they are heavily capital constrained) and b) the group is unable to post any meaningful physical collateral. Our goal then is to determine the equilibrium interest rates, where the lender must resort to a different approach for attempting to enforce the loan since there is no physical collateral or a bankruptcy code.

The following are key variables in our model. The aggregate loan for the entire group is denoted by $L$. A key assumption we make is that the group cannot engage in production in the absence of the loan. Hence, the loan is extremely attractive to the borrowers. Administrative expenses incurred by the lenders for monitoring the group is denoted by $x$ per unit time. The members of the group are subject to joint-liability, with $y$ denoting the associated expenses. This tends to reduce the value of the loan to the borrowers because of the efforts expended by the members of the group, but may improve the value of the loan due to improved performance from peer-monitoring efforts induced by joint-liability. We capture this trade-off explicitly in our model. These features have powerful ex-ante effects on how the group selects its members, and how the riskiness of the projects is chosen. By operating at the level of the group, our model will not be able to shed any insight on these ex-ante effects. As a special case, our model can also be used to examine micro-loans that are extended to individuals who are not part of a group.

We denote by $\delta(x, y)$ the proportion of wealth diverted away for consumption by the group. This is a decreasing function of the administrative expenses, $x$, incurred by the lender and the group efforts, $y$, to monitor each other. In practice, loans
have a short maturity, which we denote by $T$. The lender is assumed to employ a punishment technology to deal with ex-post defaults. In reality, such costs take the following form. The lender might be able to make the cost of entry into credit markets in future very high for borrowers who default. This is certainly a credible threat and imposes a cost if the borrowers need repeated access to credit markets. In our model, we do not consider dynamic borrowing. Hence, we may think of the cost as the present value of the difference between the rates at which the defaulting borrower should borrow in informal credit markets and the (lower) rates that would have prevailed in micro-loan markets. Also, as noted earlier, defaults in the context of group borrowing may have a significant social cost to the defaulting borrower. Anecdotal evidence suggests that exiting borrowers often pay the loans rather than defaulting to avoid such social costs. Finally, institutions such as credit bureaus can be used to reduce the incentives of borrowers to default. Punishment technology, or the credible costs associated with default, takes the form of a lump sum cost of $K$. If $K$ is too small, the borrowers will rationally take the loan and default promptly in an endogenous model of default. If it were too high, they will not borrow. It has to be high enough to induce payments of contractual obligations, which will increase the value of loans to the lender but not so large as to adversely reduce the participation of borrowers in the micro-loan programs. We will find these limits in our set up. The endogenously determined equilibrium loan rate $R$ is one of the key objects our study.

The dynamics of wealth, $C_t$, generated by the investment for borrowing group is given by the equation shown below:\footnote{We assume that $\mu$ is the expected growth rate of wealth, and $\{W_t, t \geq 0\}$ is a Brownian motion process.}

\begin{equation}
    dC_t = (\mu - \delta(x,y) - \gamma)C_t dt + \sigma C_t dW(t) \tag{3.1}
\end{equation}

It is assumed that $C(0) = L$ is the initial wealth of the borrowing group. The loan amount $L$ is specified exogenously at $t = 0$. The equilibrium borrowing rate, $R$,
is determined endogenously at $t = 0$. Group-specific risks are characterized by the constant drift and diffusive coefficients.

The specification of the process is for analytical convenience. It could be argued that it may not capture the lumpiness in income that micro-loan borrowers might face. We have solved the problem by accommodating lumpiness in output by modeling random jumps through a double exponential Poisson process, but our main results are qualitatively similar but the analytics are much more complicated. For this reason we decided to present our results for the case of the simple Brownian motion process as specified in (1).

It is useful to motivate the technology in the context of micro-loans. Micro-borrowers tend to invest their borrowing into activities such as a) livestock, b) kiosks, c) repair shops, d) paying high-interest loans to local money lenders, and e) consumption smoothing. These are typically small investments on which the rates of returns can be very high. CGAP reports estimates of rates of returns on micro-loans ranging from 40% to as high as 600%. But returns must decrease with scale, and this raises questions about the existence of a threshold scale level of micro-loans beyond which they may be less effective as a development tool unless ways are found to reduce the interest rates charged on the loans. This makes the understanding of equilibrium borrowing rates very critical.

Requiring that $L = C(0)$ makes the loan extremely valuable to the borrower because without the loan, the borrower cannot access the technology and will have a utility of zero. It is easy to introduce an endowment of $x(0)$, which will lead to the requirement that $C(0) = x(0) + L$. We have noted in our data that some borrowers have prior indebtedness when they enter the first round of borrowing in the micro-loan markets, implying that $x(0) \leq 0$. This makes the loan even more attractive to the borrowers.

Note that joint-liability has several interesting effects on budget dynamics of the borrowing group. First, the peer monitoring activity is costly to the group, and this
increases with the number of members in the group. Second, when the wealth level is low, peer monitoring, which leads to a low $\delta(x,y)$, helps to overcome potential liquidity shortages facing the group. Third, the overall risk of the project portfolio of the group represented by $\sigma$ is much lower than the risks of the projects of individual borrowers, when the group members ensure that only low risk borrowers join their group, recognizing the joint-liability feature of the loan.

Note that $\delta(x,y)$ is the "payout" or the amount diverted by the borrowing group for consumption purposes. The excess of wealth over $\delta(x,y)$, peer monitoring expenses and contractual payments is consumed in good states of the world. In bad states of the world, when the payout is less than the required payments, wealth must be liquidated to make the contractual payments.

3.3 Borrower’s Problem and Endogenous Default

The objective of the borrowing group is to maximize the discounted payoffs from the loan and select the optimal default strategy as follows. Let $\tau =\inf\{t \geq 0 : C_t \geq c^*\}$ be the first passage time of the wealth process, where $c^*$ is the borrower’s endogenously chosen optimal default trigger.

Borrower’s Problem:

$$B(C_0) = \sup_E \left[ e^{-\tau} \int_0^{\tau_T} e^{-rs}(\delta(x,y)C_s - LR)ds \right]$$

$$+ E[e^{-\tau_T} J_B(C_{\tau} - K)1_{\{\tau \leq T\}}] - Le^{-\tau_T} P(\tau > T) + E[J_B(C_T)]e^{-\tau_T} P(\tau > T)$$

where $J_B(\cdot)$ denotes the payoffs to the borrowing group upon optimally choosing to default and $r$ denotes the risk-free rates. Here we consider two possibilities. First, default leads to a lump sum punishment $K$ but the group continues to have access to technology in which the residual wealth can be invested. This is akin to saying that the tools of trade of borrowers may not be seized when default occurs. Certainly, with micro-loans the political and social costs of such an action by lenders are rather high.
One interpretation of this punishment is the following: when the borrowing group defaults, it is precluded from entering the micro-loan markets again and is forced to borrow from local money lenders at a prohibitive cost to continue to run their business. This way, the group is able to continue to have access to the technology but suffers a lump sum cost upon default. In this case, the payoff function upon default is\(^\text{10}\):

\[
J_B(C_r - K) = \frac{\delta(x = 0, y)(C_r - K)}{r + \delta(x = 0, y) + y - \mu}
\]

Alternatively, we can assume that default leads to a lack of access to the technology itself. This is a more severe punishment. The borrowing group receives a certain amount of wealth, which may be thought of as the liquidation value of their business net of punishment costs \(K\), and they must consume out of that for the rest of their lives. Given their lack of access to savings, this will constitute a more severe punishment. In this case, the payoff upon default is simply the cash flow at time of default minus the punishment cost, namely:

\[
J_B(C_r - K) = C_r - K
\]

The maximization problem of the borrower leads to the following HJB equation:

\[
0 = \max \left[ -B_t - rB + \delta(x, y)c - LR + B_c(\mu - \delta(x, y) - y)c + B_{cc} \frac{1}{2} \sigma^2 c^2 \right]
\]

(3.3)

When the wealth level of the borrowing group reaches a threshold low level \(c^*\), the group collectively defaults and receives a payoff as follows:

\[
B(c \downarrow c^*, T) = J_B(c^* - K)
\]

(3.4)

In order for the expected payoffs of the borrowing group to be finite, we need to impose

\(^{10}\) We assume that the group operates as a unit even after default and enjoy the benefits of peer monitoring after default. This can be relaxed to consider the case where default eliminates the benefits of peer monitoring.
a transversality condition. To summarize, we have specified the optimal default strategy of the borrowing group. Next, we proceed to characterize the borrower's value from taking a micro-loan of size \( L \) and defaulting optimally. The finite maturity loan is a very complicated problem and does not have a closed-form solution. We provide a very accurate approximation of the finite maturity problem, which is given in the appendix, which also explains the approximation procedure.

### 3.3.1 Equilibrium & Endogenous Borrowing Rates

We now propose a specification to describe the lender's behavior. The lender will take into account the optimal default strategy of the borrower in valuing the loan as follows. If the loan were to have a finite maturity date \( T \), the lender's problem is:

\[
D(L) = E\left[ \int_0^{rT} e^{-rs} L(R - x) ds \right] + E\left[ e^{-rT} L 1_{\{r > T\}} \right]
\]

A competitive equilibrium in this economy is an endogenously determined borrowing rate \( R \) for a loan of size \( L \) to the borrowing group such that:

1. Borrowing group's value is maximized.
2. Lender's required rate of return satisfies the fixed point requirement that \( D(L) = L \).

Note that in a competitive equilibrium the lender is not getting any surplus and the market value of the loan is exactly the amount supplied to the borrower.

---

11 We will maintain the following assumption throughout our analysis:

\[
r + \delta(x, y) + y - \mu > 0.
\]

12 The idea of the approximation is to decompose the finite maturity loan contract into two parts: 1. the no early exercise loan contract and 2. the value of the early exercise option to the borrower. This approximation is particularly good for very short and very long maturities. This approximation is particular attractive in the context of non-collateral lending as micro-loans are usually short maturity, whereas sovereign loans are usually long maturity.
We also consider a market structure in which all the surplus is extracted by the lender by setting \( B(L) = L \). We refer to this case as the monopolistic market structure, where the lender sets the interest rate \( R \) such that the borrower has no surplus. In what follows, we first characterize the competitive equilibrium.

Carrying out the integration and applying the lender's break even condition yields:

\[
R = x + r \frac{1 - e^{-rT}P(T > T)}{1 - E[e^{-rT}1_{\{T \leq T\}}] - e^{-rT}P(T > T)}
\] (3.7)

Note that the equilibrium interest rate must compensate for a) administrative expenses \( x \) borne by the lenders, b) the funding costs, which is assumed to be the risk-free rate, and c) the possibility of default by the borrowing group. Note that the probability of default will be influenced by the factors \( x, y \) and \( K \), which we will characterize in the next section.

It is easy to verify that for a perpetual loan, the equilibrium interest rate will be as follows:

\[
R = x + r \frac{1}{1 - E[e^{-rT}]}
\] (3.8)

where

\[
E[e^{-rT}] = \left( \frac{c^*}{c} \right)^{\beta_1}
\] (3.9)

and \( \beta_1 = \frac{1}{2} - \frac{\mu - \delta(x,y) - y}{\sigma^2} + \sqrt{\left(\frac{\mu - \delta(x,y) - y}{\sigma^2} - \frac{1}{2}\right)^2 + \frac{2x}{\sigma^2}} \).

The intuition of the loan rate is straightforward. The first term is the monitoring cost expended by the lender. The second term is a risk premium demanded by the lender in order to take on the risky position. The proportionality factor for risk premium demanded is \( 1 - E[e^{-rT}] \), where \( E[e^{-rT}] \) can be interpreted as the price of an Arrow-Debreu security which pays \$1 in the event of default. Note that for a 1% increase in the funding cost or the interest in the micro loan, the borrowing rate is magnified by the risk premium which is greater than 1.

The equilibrium interest rates on micro-loans is now completely characterized for given default boundary \( c^* \) and maturity \( T \), and a risk structure of default premium
can be obtained in our model. The borrowing rates under finite maturity differ from the perpetual loan rates. The key difference (in additional to different default boundaries) is in the proportionality factor for risk premium demanded. We first ignore the denominator and note that the numerator is now $1 - e^{-rT}P(T > T)$, which suggests that loan rates can be reduced. Economically, this implies the possibility of an increase in the range of equilibrium, suggesting that lower monitoring cost may be admissible in finite maturity. We explore this later and demonstrate that this is in fact the case.

Our model for the determination of loan rates differs sharply from how practitioners set the interest rate. A CGAP report indicates that the following is a standard formula for the interest rate charged in micro-finance markets:

$$R = \frac{AE + LL + CF + K - II}{1 - LL}$$ (3.10)

where, $AE$ denotes the administrative expense/monitoring cost - namely, $x$ in our model, $LL$ denotes the loan losses, $CF$ is the cost of funds - namely $r$ in our model, $K$ is the capitalization rate, and $II$ denotes the Investment Income. This equation presents a linear relationship among monitoring effort, interest rate, and joint-liability (implicit in $AE$). However, our model suggests that this relationship should be highly nonlinear. The difference is mainly due to the fact that the practitioner’s loan pricing equation ignores the fact that the loan loss rate (or defaults) depend on the borrowing rate. Furthermore, a given increase in $AE$, which is $x$ in our model, will lead to a higher loan loss rate due to the increased probability of default. The intricate dependence of these variables and the necessary risk premium are not properly reflected in the loan pricing equation of practitioners.
3.3.2 Role of Lender Monitoring & Defaults

The model delivers several implications for the role of monitoring by lenders. We now state one of the main results of our paper.

Proposition 1. In the absence of any collateral, when the borrowing group has access to technology upon default, there must be monitoring by the lender in order for there to exist an equilibrium for a perpetual debt.\(^{13}\)

The intuition behind proposition 1 is the following. When there is no monitoring by lenders, the consumption rate by the borrowing group is the same before or after default. But during the period that the loan is solvent, borrowers are forced to pay costly contractual interest payments. Immediate default allows the borrowing group to maintain its consumption and avoid paying the interest rates, which already reflect the ex-post costs of punishment. Using the parameters in the numerical illustration section, we present here the equilibrium behavior of the borrower's valuation function as a function of monitoring cost \(x\) and punishment cost \(K\) for perpetual debt. In deriving this proposition, we assume that the group still operates as a unit after default. Were this not the case, the peer monitoring benefits may still induce the group not to default even in the absence of lender monitoring.

The left panel in Figure 3.1 illustrates our proposition. Equilibrium only exists after monitoring increases up to a certain threshold level. This is due to the continuity of our problem. If monitoring decreases too much (i.e. say, below \(x = 0.2\)), there can be no equilibrium anymore. This suggests how a minimum level of lender monitoring is essential to sustain an equilibrium in micro-loan markets. Monitoring, however, can potentially be a very inefficient method for the lender to enforce the loan. We note that for each monitoring cost, there exists a \(K_{\text{max}}\) such that after this point there will not be an equilibrium anymore. To the right of this region the borrowers realize that the costs of default are too excessive for them to take the loan. We further note that

\(^{13}\) Proof is in the appendix.
$K = 0$ can be sustained as an equilibrium in our model. This is due to the fact that the borrowers' consumption is specified as an exogenous process. As long as there is monitoring, this will increase the drift of the borrower's production process. The increase in the drift is big enough to stop the borrowers from immediate default.

We now examine the situation when the maturity of the loan is very short (3 months).

**Proposition 2** In the absence of any collateral, when the borrowing group has access to technology upon default, an equilibrium exists for a finite maturity debt even in the absence of monitoring$^{14}$.

The intuition for this result is simple: the impending balloon payment of the principal accelerates the punishment costs $K$, and hence the borrower does not immediately default even in the absence of monitoring by lenders. This suggests the use of the debt maturity as a substitute for monitoring by lenders, as illustrated in the right panel of Figure 3.1. We will later show that debt maturity indeed outperforms monitoring in reducing defaults in equilibrium.

### 3.3.3 Impact of Access to Physical Collateral

In the absence of borrowing in micro-loan markets, we assume that the borrower has a utility of $\bar{U}$, which we assume to be zero for simplicity. This is the utility associated with borrowing from local money lenders, without the benefits of assortatively matched groups and peer monitoring. Once the group has access to micro-loans it is able to access the technology and create value for the borrowers. The value of the loan to the lenders is set equal to the present value of the payments promised by the borrowers, leaving the lenders with no surplus; namely $D(L) = L$. The value to the borrower is determined by the optimization problem described above. The value created by the loan to the borrower is summarized by the ratio $\frac{B(C)}{L}$. The lender can

---

$^{14}$ Proof is in the appendix.
extract some of this surplus by requiring a higher rate of return on the loan. His ability to do so will depend on the market structure within which lenders operate.

In this section we explore the case wherein the lender receives a fraction \((1 - \alpha)\) of the value conditional on default. We first modify the lender’s break-even condition to accommodate the existence of collateral upon default in our model. Let \(c^*\) be the borrower’s default strategy associated with a given loan size \(L\), then the lender’s valuation is:\(^{15}\)

\[
D(L) = E\left[\int_0^\infty e^{-rs} L(R - x) ds\right] + E[e^{-rt}(1 - \alpha)c^*] 
\]

As before, we now invoke the break even condition \(D(L) = L\) to get the equilibrium loan rate, as given by:

\[
R = x + r \frac{1 - (1 - \alpha)\frac{c^*}{L}(\frac{c^*}{c})^{\beta_1}}{1 - (\frac{c^*}{c})^{\beta_1}}
\]

where \(\beta_1 = \frac{1}{2} - \frac{\mu - \delta(x,y)-y}{\sigma^2} + \sqrt{(\frac{\mu - \delta(x,y)-y}{\sigma^2} - \frac{1}{2})^2 + 2r}.\)

Note that in the above expression, the term \(1 - (1 - \alpha)\frac{c^*}{L}(\frac{c^*}{c})^{\beta_1}\) takes account of the existence of collateral. If we set \(\alpha = 1\), we will recover our previous equation.

In our model, an increase in \(r\) typically increases the credit spreads. Also, our model would predict that in the absence of sufficient controls (as modeled by \(y, K\) and \(T\)), the defaults in micro-loans will occur sooner and spreads would be higher in micro-loans as compared to loans backed by physical collateral. To summarize,

1. Micro-loan rates are higher than loan rates backed by physical collateral.

2. Defaults in micro-loans will occur sooner than in loans backed by physical collateral, in the absence of peer monitoring, lender monitoring and punishments upon default.

---

\(^{15}\) We explore the case of perpetuity here for simplicity.
3.3.4 Competition versus Monopoly

In this section, we explore the differences in the predictions of our model in two different loan market structures: competitive and monopoly. This is done by recognizing that the borrower will take the loan as long as $B(L) \geq L$. In the limit, the lender can extract all the rents from the borrower, by setting the rate $R$ such that:

$$B(L) = L$$

The monopolistic debt lending is characterized by the interest rates $R$ such that this holds, where we denote $\hat{R}^{16}$.

Figure 3.2 explores the differences between predictions of our models as punishment cost $K$ varies. Note that, predictably, the borrower's welfare is far higher under competitive equilibrium and it increases with ex-post punishment costs $K$. The probability of default is lower and perhaps most importantly the equilibrium borrowing rates are significantly lower. Note as well that the borrowing rates are far more elastic to increases in ex-post punishment costs under a competitive market structure.

Figure 3.3 explores the differences between predictions of our models as interest rate $r$ varies. As the cost of funding $r$ increases, the default probability increases under both market structures, but the increase is far less elastic under a competitive market structure. The loan rates are much less sensitive to increases in cost of funding $r$ under the competitive market conditions. Note that under the monopolistic market structure, loan rates decline with increases in funding costs $r$. This is due to the fact that a monopolistic lender can lower the loan rate $R$ and still extract surplus from the borrower, which is not possible under the competitive structure.

Figure 3.4 explores here the differences between predictions of our models as the

---

16 Other market structure can be analyzed by setting interest rates:

$$R = \theta \hat{R} + (1 - \theta) \hat{R}$$
loan maturity $T$ varies. The important result here is that long maturity loans improve the welfare of the borrower, and keeps the interest rates low under a competitive market structure. On the other hand, the loan rates $R$ increase with maturity in a monopolistic market structure.

In the rest of the paper, we focus on a competitive loan market and examine how various contractual features influence a) loan rates, b) borrower’s welfare, and c) default probabilities. We also explore the implications of designing a loan contract wherein the contractual interest payments are tied to the flow rate of output of borrower’s technology.

### 3.3.5 Contractual Features

In order to obtain additional results, we impose the following structure on the $\delta(x, y)$ function. We need to ensure that the consumption function $\delta(x, y)$ is decreasing in monitoring and joint-liability. That is, more monitoring and peer-monitoring lead to less consumption for the borrowers. We also require that the cross derivative $\frac{\partial^2 \delta(x, y)}{\partial x \partial y} > 0$. This ensures no effect dominates each other. A particular example of such function, which we will employ throughout our analysis, is:

$$\delta(x > 0, y < y^*) = [(1 - e^{-\beta x})\dot{\delta} + e^{-\beta y} \delta] \times e^{-b \frac{y}{y^*}}$$

$$\delta(x = \infty; y) = \dot{\delta} \times e^{-b \frac{y}{y^*}}$$

$$\delta(x = 0; y = 0) = \dot{\delta}$$

In the absence of peer monitoring and lender monitoring, the borrowing group will consume an amount denoted by $\dot{\delta}$. The presence of infinite peer monitoring will lead

---

17 Specifically, we choose $\beta = -0.5$, $\dot{\delta} = 0.06$, $\delta = 0.10$, $b = 0.25$, and $y^* = 0.02$. 

to a level consumption denoted by $\delta e^{-b\frac{y}{y^*}}$. This level, $\delta$, can be thought of as the level of consumption for the borrowing group in the absence of any peer monitoring. Once the peer monitoring is put in place, its effectiveness is governed by the parameter, $b$, and the peer monitoring effort, $y$, relative to a threshold level $y^*$, which is a scaling variable. For a given effort $y$, higher the parameter $b$ is, greater is the effect of peer monitoring in controlling wasteful consumption. When $b = 0$, peer monitoring has no effect. A small $y^*$, and a high $b$, cetaris-paribus, may be thought of as a very effective peer monitoring outcome, which in reality may be due to an “assortative matching” process in which low-risk borrowers with no informational disadvantages identify other low-risk borrowers (and thereby exclude higher risk borrowers) in forming the group so that the pool in the group has low risk both in payment behavior and the riskiness of the projects undertaken by group members\textsuperscript{18}. At any effort level $y$ which is below the threshold contracting effort $y^*$, there is additional diversion of output for private consumption. A caveat is that we do not model the riskiness of the project as a function of this matching process.

With a monitoring effectiveness or efficiency parameter $\beta < 0$, lenders can in the limit approach this ideal benchmark level. In the absence of any monitoring, the existence of a punishment technology and peer monitoring through joint-liability contracting is assumed to lead to a diversion rate of $\delta > \delta$. Our specification assumes that a minimum level of monitoring $x^*$ is needed even when the group is formed under ideal conditions. In the absence of any monitoring, the existence of a punishment technology and peer monitoring through joint-liability contracting is assumed to lead to a diversion rate of $\delta > \delta$.

The delta function used in this numerical illustration is plotted in Figure 3.5. Throughout this section we have assumed the following parameters. We set the interest rate, $r$, at 3%, the drift parameter, $\mu$ at 12%, the Volatility parameter, $\sigma^2$ at 20%, the Initial Lending Amount, $L$ at 1000, the Punishment Cost, $K$, at 500 and

\textsuperscript{18} See Bannerjee, Besley and Guinnane (1994)
the maturity, $T$ at 3 months. These parameters are chosen to reflect the contractual parameters observed in practice. The average size of micro-loans are three months and are relatively small.

Figure 3.6 investigates the range of equilibrium and how different amount of monitoring cost and joint liability affects equilibrium. Note that with short maturity, excessive monitoring reduces the welfare of the borrower. First, default probability is increasing in monitoring. This is a consequence of the high loan rates charged by the lenders to compensate for the increased monitoring costs. With short maturity, the borrowers are now faced with higher repayment rates, and since the technology may not be able to produce enough cash flow for repayments in such short maturity, the borrowers are forced to default. Although the lender can transfer all the cost incurred to the borrower, the lender still has to suffer from lower repayment probabilities. This is due to the feedback effect of increased probability of default induced by the higher borrowing rates in our equilibrium analysis, which is absent in the practitioner model. Hence, practitioners believe lender monitoring can lead to lower defaults, whereas it is not necessarily the case in our model. This suggests that finite maturity is an useful substitute for monitoring. It reduces default probability while keeping loan rates feasible for the borrowers. Joint liability, on the other hand, serves as another substitute for monitoring. Default probability is concave in joint liability, suggesting that when joint liability is low, default probability increases. However, once the borrowers exert too much joint liability, they do not concentrate in the production process enough and causing more defaults in equilibrium. This effect translates to a concave borrower’s valuation function. Loan rates are decreasing in joint liability as a very slow rate.

Figure 3.7 investigates the dynamics of the equilibrium quantities in terms of punishment cost $K$ and joint liability $y$. We will analyze values joint-liability for $y = 0\%$ (solid line), $y = 0.5\%$ (dashed line), and $y = 1\%$ (dotted line). We will fix $x = 30\%$. Several aspects of these pictures are worthy of additional discussion: first,
note that a higher level of peer monitoring leads to a lower probability of default at all levels of $K$. Second, at higher levels of peer monitoring as proxied by the variable $y$, the range of punishment costs $K$ conditional on default is much higher: in other words, the default probability is a much flatter function of $K$ at higher levels of $y$. Although the loan rates are lower with higher levels of $y$, the borrower's value function is declining in $y$ as the borrowing group is forced to put in the peer monitoring effort.

Figure 3.8 explores how equilibrium changes as a function of punishment cost as we increase monitoring cost for $x = 20\%$ (solid line), $x = 25\%$ (dashed line), and $x = 30\%$ (dotted line). We will fix $y = 0.5\%$. As the monitoring effort $x$ by the lenders increases, the borrowing costs increase as well. We note that monitoring increases default probability while raising loan rates. It also lowers borrower's value due to the increased defaults. Hence, in short maturity debt contracts, monitoring do not play an effective role.

Figure 3.9 explores the relationship between the borrower's and the lender's actual cost of lending. As before, we focus on the cases where an equilibrium exists. Namely, we have $x=20\%$ (solid line), $x=25\%$ (dashed line), and $x=30\%$ (dotted line). Default probability and loan rates are increasing in both the monitoring expense and the interest rate. Loan rate is increasing in interest rate, as the lender demands a higher risk premium to compensate for his increasing opportunity cost. Furthermore, higher administrative cost induces a higher equilibrium loan rate. As a result of the higher loan rates demanded by the lender, default probability is increasing in loan rates. These effects force the borrower's value to be decreasing in monitoring expense and interest rate. The intuition behind these results is that as interest rate goes up, loan rates increase much more significantly. This suggests that the cost of funding for the lender may be a key to the determination of micro-loan interest rates.

Figure 3.10 explores the term-structure implied by our model. The lines plot the term structure of normalized borrower's value, default probability, and equilibrium loan rates for $L = 1000$ (solid line), $L = 5000$ (dashed line), and $L = 10000$ (dotted
line. Note that for fixed punishment cost $K$, our model requires large loans and long maturities. Small loans on the other hand require short maturities. Note that as $K$ increases from 500 to 750, the short maturities become admissible for loan size of 1000. This suggests the use of punishment cost as a very powerful device for enlarging the range of equilibriums. Finally, as the loan size increases, the probability of default increases and the loan rates dramatically increase unless the maturity of the loans are increased.

We now present the trade off between monitoring and debt maturity and show the choice of debt maturity can be used as a substitute to monitoring. We focus on $x=20\%$ (solid line), $x=25\%$ (dashed line), and $x=30\%$ (dotted line). Figure 3.11 clearly show that debt maturity can be used to control defaults in equilibrium. Short maturity loans almost never default while long maturity loans are more inclined to default. On the other hand, increasing monitoring simply increases loan rates and default probability. This suggests that the use of debt maturity as a contracting device for the lenders instead of monitoring.

In summary, a prevalent effect in our model is that default probability and loan rates increase in monitoring. This suggests that when practitioners decide their micro-lending business strategy, monitoring costs should be given extra concern. If the group is properly formed (i.e., for a reasonable joint liability $y$), and a fixed opportunity cost $r$, micro-finance institutions may achieve a lower default rate by reducing monitoring. Furthermore, the micro-finance institution may substitute monitoring using maturity $T$ of debt as a tool. Finally, punishment cost $K$ can also be used to alter equilibrium risk structure.

### 3.3.6 Fixed Rate Loan vs. Floating Rate Loan

In this section, we explore the welfare consequences of designing the loan contract so that the interest payments are proportional to the flow rate of output from the technology of the borrowing group. We refer to this case as a floating rate loan. Es-
sentially, with a floating rate loan, lenders permit borrowers to make higher payments in "good states" to insure against lower payments in "bad states" but keep the present value of payments equal to the loan amount $L$ in a competitive market. A complete solution of the problem is given in the appendix. The contractual coupon payments at each instant are $\hat{\beta}C_t$, and the lender still enforces the break-even condition as before. Analysis of this case leads to the following important result:

- $c^{\text{Fixed}}_t \geq c^{\text{Floating}}_t$, meaning that in fixed rate specification, the borrower defaults earlier. This leads to the implication that the default probabilities are lower if the loan contract permits a variable interest payment schedule that is linked to the flow rate of output of the borrowing group.

In the floating rate loan specification, default decision is independent of the interest repayment rate, which is labeled as $\hat{\beta}$ in the solution given in the appendix. The intuition is that the lenders now bear all the risk with the borrowers. The range of admissible equilibrium is greater under fixed rate specification than under floating rate specification. Given that the lenders know the default strategy of the borrowers, lenders do not lend as much regardless of the rate they charge since they know it does not affect the borrower's decision.

### 3.4 Conclusion

We have presented a simple model of lending without collateral. The lender attempts to enforce the contract by relying on three things: a) monitoring to reduce the diversion of resources by the borrower from productive uses, b) peer monitoring by lending to a group, which is jointly liable for the fulfillment of the contractual provisions, and c) a punishment technology that imposes a finite cost on defaulting group of borrowers. We show that peer monitoring combined with a limited amount of monitoring by lenders is sufficient to reduce default probability to acceptable levels, so long as there is a credible punishment cost. Excessive monitoring by lenders increases the
cost of borrowing and this might lead to non-participation by borrowers. As the loan size increases, we show that the probability of default increases, and the loan rates dramatically increase, unless the maturity of the loans is increased.

We have extended our analysis to examine situations where the borrowers face low frequency jump risks. Episodes such as heavy monsoons or health epidemics could have dire consequences for borrowers in this market. Predictably, we found initial loan rates to be too prohibitive in the presence of adverse jump risks. An important limitation of our work is that we do not examine repeated borrowing and the discipline that may impose on the borrowing group.

### 3.5 Appendix

In all our derivations, let \( \tau = \inf \{ t > 0 : C_t > c^* \} \) be the first passage time of the cash flow process.

#### 3.5.1 Fixed Rate Debt Contract

**GBM Perpetual Loan Contract**

Since most structural corporate debt models assume perpetual debt, we present here the borrower’s valuation with perpetual loans:

\[
B(c) = B(\alpha c - K)
\]  

(3.13)

for \( c \geq c^* \):

\[
B(c) = A_1 \left( \frac{c^*}{c} \right)^{\rho_1} + A_3 c + A_4
\]

(3.14)

where

\[
A_1 = (\alpha B - A_3)c^* - (BK + A_4)
\]

\[
A_3 = \frac{\delta(x, y)}{r + \delta(x, y) + y - \mu}
\]
and \( B \) is defined above, which can take on two values depending whether we allow for access to technology or not upon default. The interpretation of the above formula straightforward. The first term is the risk neutral expectation of the investment technology net of punishment cost and default risk. The second term is the value of the technology up on default. The last term is the present value of the total cost exerted by the borrower.

Proof Let \( x = \log(C) \) and assume:

\[
B(C) = A_1 e^{-x\beta_1} + A_3 e^x + A_4
\]

We will suppress the dependencies of \( \delta \) for convenience. Substituting this into the HJB equation yields:

\[
0 = A_1 e^{-x\beta_1} (+r - (\mu - \delta - y - \frac{1}{2}\sigma^2)(-\beta_1) + \frac{1}{2}\sigma^2\beta_1^2) \\
+ e^x((r - \mu - \delta - y)A_3 + \delta) \\
-LR - rA_4
\]

Now, by the technique of matching the coefficient, we can solve for \( \beta_1, A_3 \) and \( A_4 \) in closed form. \( A_1 \) is obtained via the Principal of Continuity:

\[
Bae^{x_0} - BK = A_1 e^{-\beta_1 x_0} + A_3 e^{x_0} + A_4
\]
where $B$ is defined in the paper, referring to different values according to whether the borrower has access to technology or not after default.

Finally, the Principal of Smoothing Pasting gives the optimal default boundary:

$$B ae^{x_0} = -A_1 \beta_1 e^{-\beta_1 x_0} + A_3 e^{x_0}$$

We have 5 equations and 5 unknowns, giving us an identified system to solve for: $\beta_1$, $A_1$, $A_3$, $A_4$ and $x_0$. Now recognizing that $c^* = e^{x_0}$, the proof is complete.

GBM Finite Maturity Loan

The borrower’s value function is given by: For $c < c^*$:

$$B(c, T) = B(c - K)$$

(3.15)

for $C \geq c^*$:

$$B(C_0, T) = EuB(C_0, T) + A_1 \left( \frac{c^*}{c} \right)^{\beta_1} + A_3 c + A_4$$

(3.16)

where

$$A_1 = (B - A_3 - Q)c^* - (BK + A_4)$$

$$A_3 = \frac{\delta(x, y)/z}{r/z + \delta(x, y) + y - \mu}$$

$$A_4 = \frac{LR}{r}$$

$$c_{approx} = \frac{(BK + A4)\beta_1}{(B - A3 - Q)(1 + \beta_1)}$$

$$Q = B e^{(\mu - \delta(x, y) - y)T}$$

$$z = 1 - e^{-rT}$$

$$\beta_1 = \frac{1}{2} \frac{\mu - \delta(x, y) - y}{\sigma^2} + \sqrt{\left(\frac{\mu - \delta(x, y) - y}{\sigma^2} - \frac{1}{2}\right)^2 + \frac{2r/z}{\sigma^2}}$$

The constant $B$ is again chosen according to whether there is access to technology or not up on default. $EuB(C_0, T)$ represents the European version of the same debt.
contract, and is given by:

\[ EuB(C_0) = -Le^{-rT} + e^{-rT}E[J_B(C_T)] \]
\[ = -Le^{-rT} + BC_0 e^{(\mu-r-\delta(x,y)-y)T} \]
\[ = -Le^{-rT} + QC_0 \]

where \( B = \frac{\delta(x=0,y)}{r+\delta(x=0,y)+y-\mu} \).

Note that \( c_{approx}^* \) converges to \( c^* \) and the finite maturity approximation of the borrower’s value converges to the perpetuity function as \( T \) goes to infinity, verifying the accuracy of our approximation scheme.

From the finite maturity approximation, we see that even with \( x = 0 \), there can be an equilibrium - a main difference between the finite maturity and the perpetual conclusion. The intuition is that finite maturity itself is also an additional tool both the lender and the borrower can use to screen out unwanted loan contracts. Although the consumption rate is still the same before and after default, with the additional constraint of finite maturity, the borrowers will choose not to immediate default as long as the amount of interest paid is less than the punishment cost of defaulting.

**Proof** In the following, we will suppress the dependencies of \( \delta \) for convenience.

The borrower’s value has to satisfy the following HJB equation: For \( c \geq c^* \):

\[ 0 = \max \left[ -B_t - rB + \delta c - LR + B_c(\mu - \delta - y)c + B_{cc} \frac{1}{2} \sigma^2 c^2 \right] \]

(3.17)

and

\[ B(c,T) = J_B(c - K) \]

(3.18)

By Feymann-Kac, we know that \( EuB(c,T) \) solves the following partial differential equation for all \( c \):

\[ 0 = \max \left[ -EuB_t - rEuB + \delta c - LR + EuB_c(\mu - \delta - y)c + EuB_{cc} \frac{1}{2} \sigma^2 c^2 \right] \]

(3.19)
Hence, the early exercise premium \( \epsilon(c, T) \) must satisfy: For \( c \geq c^* \):

\[
-r\epsilon + (\mu - \delta - y)\epsilon_c + \frac{1}{2}\sigma^2\epsilon_{cc} + \delta(x, y)c - LR = 0
\]  
(3.20)

and for \( c \leq c^* \):

\[
\epsilon(c, T) = J_B(c - K) - EuB(c, T)
\]  
(3.21)

Now by letting \( z = 1 - e^{-rt} \) and \( g(c, z) = \frac{\epsilon(c, T)}{z} \). It is easy to see that: \( z_t = re^{-rt} \), \( \epsilon_x = zg_x, \epsilon_{xx} = zg_{xx} \) and \( \epsilon_t = z_t g + z g_z z_t \). Substitute this into the HJB and divide by \( z \), the HJB becomes: For \( c \geq c^* \):

\[
-r(1 - z)g_x - \frac{r}{z}g + \frac{1}{2}\sigma^2 zg_{xx} + (\mu - \delta - y)zg_c + \frac{\delta}{c} c - LR = 0
\]  
(3.22)

We will assume that \((1 - z)g_z = 0\) for the approximation. This approximation becomes very accurate for very short and very long maturity.

We are now in the position to solve this equation. We recognize this as the Euler's equation and hence assume:

\[
\epsilon(c) = A_1 \left( c^* \right)^{\beta_1} A_3 c + A_4
\]

By the method of matching coefficients, we get the following equations:

\[
-r/z + (\mu - \delta - y - \frac{1}{2}\sigma^2)(-\beta_1) + \frac{1}{2}\sigma^2\beta_1^2 = 0
\]

\[
(-\frac{r}{z} + \mu - \delta(x, y) - y)A_3 + \frac{\delta}{z} = 0
\]

\[
rA_4 + LR = 0
\]

Imposing continuity at the boundary:

\[
\epsilon(c^*, T) = B(c^* - K) - EuB(c^*, T)
\]  
(3.23)
allows us to solve for the coefficient $A_1$ as a function of optimal default boundary $c^*$. Imposing the principal of smooth fit:

$$e'(c^*) = Bc^* - \frac{\partial}{\partial x} EuB(c^*, T)|_{x=x_0}$$  \hspace{1cm} (3.24)

gives the optimal default boundary.

We have a system of 5 equations and 5 unknowns and hence all the variables are identified. This completes the proof.

\subsection*{3.5.2 Proof of Proposition 1}

\textbf{Proof} Step 1.

First, we will show that the value of continuation is always less than value of defaulting immediately, when $x = 0$.

Case 1. $c^* > L$

There is no equilibrium by definition.

Case 2. $c^* < L$

Suppose $c^* = c_{\text{bar}} < L$, then by transversality condition, the borrower’s value function is well defined and finite. Hence, $c^*$ must satisfy the equation for the optimal default boundary:

$$c^* = \frac{(BK + A4)^{\frac{\beta_1 \beta_2}{n_2}}}{(B - A3)^{\frac{1+\beta_1}{1+\beta_2}}}$$

Note that the LHS is finite by assumption. The numerator of RHS depends on $R$, which we can calculate given default boundary using (3.8). However, the denominator is 0 and hence $c^*$ is undefined, contradicting the fact that transversality condition (3.5) guarantees a well-defined value’s function.

Case 3. $c^* = L$

When $c^* = L$, for any triplet $(K, x, y)$ the value of continuation equals the value of immediate default by the \textit{Principle of Smooth Fit} (i.e. the value function is continu-
ous at $c^*$).

Step 2.
Let us now show that $c^* = L$ cannot be an equilibrium for any punishment cost $K$.

Case 1. $K < L$
If the punishment cost is less than initial loan amount, the borrower's optimal strategy is to default immediately. However, the lender knows it and hence she won't lend.

Case 2. $K > L$
If the punishment cost is higher than initial loan amount, the borrower's value function will always be negative since $J_B(L) < 0$. Hence, the borrower is better off not borrowing.

Case 3. $K = L$
The borrower has a value function exactly 0. Hence, she is indifferent between lending and borrowing.

3.5.3 Collateralized versus Micro-loan Rates

Proof We note that in Leland(1994)'s model, we can rewrite the default boundary for a collateralized loan as:

$$c_{col}^* = \frac{C}{r} \frac{\beta_{corp}}{1 + \beta_{corp}}$$

Similarly, in the absence of access to technology upon default and the drift of the technology process restricted to $r$ for the existence of a martingale measure, we have $A_3 = 1$ and $B = 1$. This gives:

$$c^* = \frac{LR \beta_3}{r(1 + \beta_3)} \frac{1}{1 - (1 - \alpha)}$$
The first part of the proposition follows immediately once we note that $\alpha \geq 1$ is a natural assumption. After rearranging for $R$, the same analysis leads to the second result.

3.5.4 Numerical Procedure

This section documents the numerical procedure in solving for the equilibrium $(R, c^*)$. We solve for our equilibrium as follows:

1. In the equation for $c^*$, we plug in the equation for equilibrium loan rate $R$, which depends on $c^*$ as well. Note that the equations for $R$ in the finite maturity case involves the terms $P(\tau > T)$ and $E[e^{-\tau r}1_{\{r \leq T\}}]$, which we use the method of Laplace transforms to obtain. Specifically, we apply the Gaver-Stehfest inversion algorithm to the Laplace transforms of $P(\tau > T)$ and $E[e^{-\tau r}1_{\{r \leq T\}}]$.

2. We numerically vary $c^*$ until the fixed point equation for $c^*$ is satisfied.

3. We then use the solution for $c^*$ to get our equilibrium loan rate $R$.

4. If $c^*$ is within the admissible range $(0, L)$, then we check whether the borrower's valuation function is positive. If so, we have an equilibrium. Otherwise, there is no equilibrium.

3.5.5 Floating Rate Debt Contract

GBM Perpetual Loan Contract

We present here the borrower's problem with floating rate, by which we mean that the instantaneous repayment for the borrower is $\beta C_t dt$. The HJB equation is given by:

$$0 = \max \left[ -B_t - r B + (\delta - \beta)c + B_c(\mu - \delta - \gamma)c + B_{cc} \frac{1}{2} \sigma^2 c^2 \right]$$

(3.25)
The solution of this equation with $B_t = 0$ is given by:

$$B(c) = B(c - K)$$  \hspace{1cm} (3.26)

for $c \geq c^*$:

$$B(c) = A_1\left(\frac{c^*}{c}\right)^{\beta_1} + A_3c$$  \hspace{1cm} (3.27)

where

\begin{align*}
A_1 &= (\alpha B - A_3)c^* - BK \\
A_3 &= \frac{\delta(x, y) - \beta}{r + \delta(x, y) + y - \mu} \\
c^* &= \frac{BK\beta_1}{(\alpha B - A_3)(1 + \beta_1)} \\
B &= \frac{\delta(x = 0, y)}{r + \delta(x = 0, y) + y - \mu} \\
\beta_1 &= \frac{1}{2} - \frac{\mu - \delta(x, y) - y}{\sigma^2} + \sqrt{\left(\frac{\mu - \delta(x, y) - y}{\sigma^2} - \frac{1}{2}\right)^2 + \frac{2r}{\sigma^2}}
\end{align*}

and $B$ is defined above, which can take on two values depending whether we allow for access to technology or not upon default.

**GBM Finite Maturity Loan**

The borrower's value function is given by: For $c \leq c^*$:

$$B(c, T) = B(c - K)$$  \hspace{1cm} (3.28)

for $C \geq c^*$:

$$B(c, T) = EuB(c, T) + A_1\left(\frac{c^*}{c}\right)^{\beta_1} + A_3c$$  \hspace{1cm} (3.29)

where

\begin{align*}
A_1 &= (B - A_3 - Q)c^* - BK
\end{align*}
\[
A_3 = \frac{(\delta(x, y) - \beta)/z}{r/z + \delta(x, y) + y - \mu} \\
\beta_1 = \frac{(B - A3 - Q)(1 + \beta_1)}{BK\beta_1} \\
Q = B_0(\mu - r - \delta(x, y) - y)^T \\
z = 1 - e^{-rT} \\
\beta_1 = \frac{1}{2} - \frac{\mu - \delta(x, y) - y}{\sigma^2} + \sqrt{\left(\frac{\mu - \delta(x, y) - y}{\sigma^2} - \frac{1}{2}\right)^2 + \frac{2r/z}{\sigma^2}}
\]

The constant \( B \) is again chosen according to whether there is access to technology or not up on default. \( EuB(c, T) \) represents the European version of the same debt contract, and is in the finite maturity section for the fixed rate section above.

Some Discussions

1. Note that \( c^*_{\text{Fixed}} \geq c^*_{\text{Floating}} \), meaning that in fixed rate specification, the borrower defaults earlier.

2. In the floating rate specification, default decision is independent of the interest repayment rate \( \beta \). The intuition is that the lenders now bear all the risk with the borrowers.

3. The range of admissible equilibrium is greater under fixed rate specification than under floating rate specification. Given that the lenders know the default strategy of the borrowers, lenders do not lend as much regardless of the rate they charge since they know it does not affect the borrower’s decision.
<table>
<thead>
<tr>
<th>REGION</th>
<th>Gross Loan Portfolio Size in US $</th>
<th>Active Borrowers Number</th>
<th>Per Capita Loan Size</th>
<th>Loan Rates</th>
<th>Write-off Ratio</th>
<th>PAR ≥ 30 Days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa</td>
<td>1,010,088,380 (12%)</td>
<td>3,154,502 (11%)</td>
<td>320.21</td>
<td>39.84%</td>
<td>3.32%</td>
<td>9.15%</td>
</tr>
<tr>
<td>East Asia</td>
<td>1,983,635,418 (23%)</td>
<td>4,103,326 (15%)</td>
<td>483.42</td>
<td>39.10%</td>
<td>4.12%</td>
<td>7.83%</td>
</tr>
<tr>
<td>East Europe</td>
<td>1,449,653,047 (17%)</td>
<td>789,396 (3%)</td>
<td>1,835.15</td>
<td>28.86%</td>
<td>0.60%</td>
<td>1.78%</td>
</tr>
<tr>
<td>Latin America</td>
<td>2,740,536,803 (31%)</td>
<td>3,231,062 (11%)</td>
<td>848.18</td>
<td>35.82%</td>
<td>2.52%</td>
<td>6.62%</td>
</tr>
<tr>
<td>Middle East</td>
<td>208,032,901 (2%)</td>
<td>794,083 (3%)</td>
<td>261.98</td>
<td>34.61%</td>
<td>0.15%</td>
<td>2.60%</td>
</tr>
<tr>
<td>South Asia</td>
<td>1,327,858,980 (15%)</td>
<td>16,057,919 (57%)</td>
<td>82.69</td>
<td>20.29%</td>
<td>0.74%</td>
<td>10.44%</td>
</tr>
<tr>
<td>TOTAL:</td>
<td>8,719,805,529 (100%)</td>
<td>28,130,828 (100%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AVERAGE:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3.1: A decomposition of the Micro-Loan Market by regions. The data is collected using the MIX database for the year 2003. The numbers in the parentheses under Gross Loan Portfolio and Active Borrowers indicate their proportions in the total market. Per Capita Loan Size is reported in US dollars.

<table>
<thead>
<tr>
<th>Lender</th>
<th>Gross Loan Portfolio Size in US $</th>
<th>Active Borrowers Number</th>
<th>Per Capita Loan Size</th>
<th>Loan Rates</th>
<th>Write-off Ratio</th>
<th>PAR ≥ 30 Days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Banks (47)</td>
<td>4,502,900,041 (51.6%)</td>
<td>8,853,649 (31.5%)</td>
<td>509</td>
<td>24.39%</td>
<td>0.51%</td>
<td>9.61%</td>
</tr>
<tr>
<td>Cooperatives and Credit (98) Unions</td>
<td>701,205,569 (8.0%)</td>
<td>805,018 (2.9%)</td>
<td>871</td>
<td>30.92%</td>
<td>3.29%</td>
<td>8.58%</td>
</tr>
<tr>
<td>Non-bank financial (49) Institutions</td>
<td>1,627,919,395 (18.7%)</td>
<td>4,827,168 (17.2%)</td>
<td>132</td>
<td>37.80%</td>
<td>0.83%</td>
<td>4.71%</td>
</tr>
<tr>
<td>NGO (286)</td>
<td>1,509,926,890 (17.3%)</td>
<td>13,074,367 (46.5%)</td>
<td>115</td>
<td>35.75%</td>
<td>2.48%</td>
<td>7.52%</td>
</tr>
<tr>
<td>Others (32)</td>
<td>346,857,110 (4.0%)</td>
<td>425,679 (1.5%)</td>
<td>815</td>
<td>37.54%</td>
<td>1.51%</td>
<td>4.26%</td>
</tr>
<tr>
<td>Rural bank (15)</td>
<td>30,996,524 (0.4%)</td>
<td>144,947 (0.5%)</td>
<td>214</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL:</td>
<td>8,719,805,529 (100%)</td>
<td>28,130,828 (100%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AVERAGE:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3.2: A decomposition of the Micro-Loan Market by lenders. The data is collected using the MIX database for the year 2003. The numbers in the parentheses under the column labeled Lender is the number of voluntary reporting institutions for each category of lenders. The numbers the parentheses under Gross Loan Portfolio and Active Borrowers indicate their proportions in the total market. Per Capita Loan Size is reported in US dollars.
Figure 3.1: The left panel displays the borrower’s value as a function of monitoring cost ($x$) and punishment cost ($K$) in the case of a perpetual loan. The right panel displays the borrower’s value as a function of monitoring cost ($x$) and punishment cost ($K$) for finite maturity $T = 3$ months.

Figure 3.2: This set of figures analyze the effect of two market structures: monopolistic and competitive lending. The equilibrium behaviors as ex-post punishment cost ($K$) varies are displayed. The left panel displays the normalized borrower’s value. The middle panel displays the default probability. The right panel displays the equilibrium loan rates.
Figure 3.3: This set of figures analyze the effect of two market structures: monopolistic and competitive lending. The equilibrium behaviors as lender’s cost of funding \( r \) varies are displayed. The left panel displays the normalized borrower’s value. The middle panel displays the default probability. The right panel displays the equilibrium loan rates.

Figure 3.4: This set of figures analyze the effect of two market structures: monopolistic and competitive lending. The equilibrium behaviors as maturity \( T \) varies are displayed. The left panel displays the normalized borrower’s value. The middle panel displays the default probability. The right panel displays the equilibrium loan rates.

Figure 3.5: This illustrates our choice of the borrowers’ consumption ratio, as captured by the \( \delta(x, y) \) function, as a function of monitoring cost \( x \) and joint liability \( y \).
Figure 3.6: Equilibrium behavior as a function of administrative cost (\(x\)) and joint-liability (\(y\)) when there is access to technology upon default. The left panel displays the borrower's value normalized by the initial borrowing amount. The middle panel displays default probability. The right panel displays equilibrium loan rates.

Figure 3.7: Equilibrium behavior as joint-liability (\(y\)) increases. The values analyzed are for joint-liability \(y = 0\%\) (solid line), \(y = 0.5\%\) (dashed line), and \(y = 1\%\) (dotted line). Administrative cost is fixed at \(x = 30\%\). The left panel displays the borrower's value normalized by the initial borrowing amount. The middle panel displays default probability. The right panel displays equilibrium loan rates.

Figure 3.8: Equilibrium behavior as administrative cost (\(x\)) increases. The values examined are for administrative cost \(x = 20\%\) (solid line), \(x = 25\%\) (dashed line), and \(x = 30\%\) (dotted line). Joint-liability is fixed at \(y = 0.5\%\). The left panel displays the borrower's value normalized by the initial borrowing amount. The middle panel displays default probability. The right panel displays equilibrium loan rates.
Figure 3.9: Equilibrium behavior as lender's cost of funding (r) increases. We fix administrative costs to be x=20% (solid line), x=25% (dashed line), and x=30% (dotted line). The left panel displays the borrower's value normalized by the initial borrowing amount. The middle panel displays default probability. The right panel displays equilibrium loan rates.

Figure 3.10: Equilibrium behavior as the maturity $T$ increases. The monitoring cost examined are $x = 20\%$ (solid line), $x = 25\%$ (dashed line), and $x = 30\%$ (dotted line). We fix $y = 0.5\%$. The left panel displays the borrower's value normalized by the initial borrowing amount. The middle panel displays default probability. The right panel displays equilibrium loan rates.

Figure 3.11: Equilibrium behavior as maturity $T$ increases. We focus on $x=20\%$ (solid line), $x=25\%$ (dashed line), and $x=30\%$ (dotted line). We fix $y = 0.5\%$. The left panel displays the borrower's value normalized by the initial borrowing amount. The middle panel displays default probability. The right panel displays equilibrium loan rates.
Bibliography


Rytchkov, O., 2008, "Expected Returns on Value, Growth, and HML," working paper, University of Texas at Austin.
