University of Alberta

Modeling Non-linearity in Asset Returns

by

©

Wing Hong Chan

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of **Doctor of Philosophy**.

Department of Economics

Edmonton, Alberta
Fall 2002

0-612-81171-9

University of Alberta

Library Release Form

Name of Author: Wing Hong Chan

Title of Thesis: Modeling Non-linearity in Asset Returns

Degree: Doctor of Philosophy

Year this Degree Granted: 2002

Wing Chan

Wing Hong Chan
5/F, 22, The First Street
Sai Ying Pun
Hong Kong

Date: June 25, 2002

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **Modeling Non-linearity in Asset Returns** submitted by Wing Hong Chan in partial fulfillment of the requirements for the degree of **Doctor of Philosophy**.

Purkan Dutt

Dr. J. Maheu (Co-Supervisor)

Denise Young

Dr. D. Young (Co-Supervisor)

A. Buse

Dr. A. Buse

Dr. S. Vigeant

Dr. A. Kaul

P.D.

Dr. J. Knight (External Reader)

Date: June 18, 2002.

Abstract

The first essay develops a new conditional jump model to study jump dynamics in stock market returns. We propose a simple filter to infer *ex post* the distribution of jumps. This permits construction of the *shock* affecting the time t conditional jump intensity, and is the main input into an autoregressive conditional jump intensity model. The model allows the conditional jump intensity to be time-varying and follows an approximate ARMA form. Daily stock returns are analyzed using the jump model coupled with a GARCH specification of volatility. We find significant time-variation in the conditional jump intensity and evidence of time-variation in the jump size distribution. The conditional jump dynamics contribute to a good in-sample and out-of-sample fit to stock market volatility, and capture the rally often observed in equity markets following a significant downturn.

The second essay develops a new bivariate jump model to study jump dynamics in foreign exchange returns. The model extends a bivariate GARCH parameterization to include a correlated jump process. The conditional covariance matrix has the BEKK structure, while the bivariate jumps are governed by a Correlated Bivariate Poisson (CBP) function. Using daily data we find evidence of both independent currency-specific jumps and jumps common to both exchange rates. The essay concludes by investigating a time-varying structure for the arrival of jumps that relaxes the assumption of a constant and bounded jump correlation imposed by the CBP function.

The third essay extends the recent empirical literature on risk-adjusted Hotelling Rules for exhaustible resources by allowing time-varying risk premia in resource asset

returns. Monthly metal prices for four metals (lead, copper, silver, and zinc) are used to estimate a Bivariate GARCH-M model derived from the risk-adjusted Hotelling rule and the consumption-based Intertemporal Asset Pricing Model (IAPM) of resource asset returns. The model is also augmented with the Correlated Bivariate Poisson jump component to study the effect of market crashes on resource returns. The results basically reject the hypothesis that a time-varying risk premium provide a key component in modeling nonrenewable resource asset returns.

Acknowledgements

I would like to first thank God for everything he has given me. Erica is always by my side and shares good and bad times. She is a very understanding wife and she means everything to me. Without her, no food tastes good, no subject is worth studying, and no degree is worth getting. Her moral support is very important to the completion of this thesis.

There are three people that I have no words to describe how lucky I am to have met them and working with them for the last six years. I thankfully acknowledge the help from my two co-supervisors, John Maheu and Denise Young. Without their ideas and help, I would never made it through these six years. I will treasure the valuable knowledge that I have learned from them and pass it onto my students.

Adolf Buse, who showed me the importance of being a good teacher, can not be forgotten. He also helped me to take the first step toward understanding Econometrics. He made me realize the fact that you can teach anyone Econometrics as long as you are willing and patient. A dumb student like me provides one solid evidence.

I also like to remember the help from the department staff, Audrey Jackson, Charlene Hill, Leslie Wayne, Louise Edwards, and Margaret Howell. They make me feel like part of a big family where everyone is caring for each other and I can count on their help and support. Knowing that I have to leave them brings great sadness.

I am also indebted to Melville McMillan for the opportunity of working with him in numerous projects over the last few years. His Christmas dinner is always the highlight of a year for foreign students like Erica and me. Finally, I wish to thank Aditya Kaul, John Knight, and Stéphane Vigeant for helpful comments.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Modeling financial asset returns is a challenging task. With every day, new data points are available and there are frequent discoveries of (often puzzling) empirical characteristics. The goals in this dissertation are twofold. The first objective is a quest for a better understanding of market dynamics. Improved knowledge regarding the elements determining market movements may increase the ability of the government to provide stability in the financial system. Furthermore, the private sector can benefit from a less risky investment environment. The second objective is to provide accurate forecasts of the financial series, which has important implications for risk management, optimal hedging, and production planning related to various types of financial instruments.

Classical linear models, although often well-suited to the understanding of relationships between economic variables, are ill-suited when used as a forecasting tool with respect to the financial markets. This results mainly from the complexity of financial markets and the speculative nature of the system. A shift of focus from linear modeling to nonlinear modeling has occurred in the literature and well-developed nonlinear models have been presented in various contexts with great success. A best nonlinear model for the financial markets is difficult to identify. However, there is a common belief that in general these models perform well in a variety of situations with the most appropriate model depending on the particular series being studied, the time period selected, and the sample size (see Gençay, Ballocchi, Daorogna, Olsen, and Pictet (2002), Gençay, Daorogna, Olsen, and Pictet (2002), Daorogna, Gençay, Müller, Olsen, and Pictet (2001)).

1.1 Motivations

The unifying theme of this thesis relates to the possibility of identifying predictable components of asset returns. There are two main focuses pertaining to market crashes and risk premia. A question immediately arises, as no crashes should have occurred if they are predictable. The results of this thesis will allow us to argue that more information can be gained by carefully examining the history of dramatic market movements than by simply naming them crashes with total ignorance. It may be impossible to accurately predict when the next "1987" crash will come, but predictable components can be found with careful investigation.

Market crashes from hereon forward will be referred to as jumps that should be carefully defined not only to include the severe downturns of the market, but any unusually large movements following the entry of unanticipated news into the market. Looking at jumps from a statistical perspective, consider a data series that is normally distributed and, therefore, the chance of the occurence of an observation beyond three standard deviations is very low. For example, in the case of Canadian dollar, given that it has experienced returns of 0.01% per day on average for the last ten years with a standard deviation equal to 0.25%, a 1% jump (or a half-cent jump in terms of level) on a single day would be an unlikely event. Historically, though it happens so often that a normal distribution can not be used to explain this phenomenon. We denote these abrupt changes as crashes or jumps in this thesis.

A few questions motivate the line of research in this thesis. These include: "Is there a systematic pattern in the arrival rate of market jumps?", "Do the magnitudes of these jumps have implications for future jumps?", "Does the likelihood of jumps remain constant over time?", "What can be learned from the past that can be used to infer currency or stock market jumps in the future?", "Will market jumps spill over from one market to the next?". This thesis attempts to answer these questions in a analytical and technical framework and presents nonlinear models that are capable of answering these questions in an hypothesis testing framework as well as through direct incorporation of stylized facts. The proposed models will be applied to three different markets: stock markets, foreign exchange markets, and commodity (metal) markets.

Risk premia are the focus of the most important areas in the literature due to the potential profit implications. The difficulty of modeling such a component comes from the fact that the sources of risk premia are usually unknown. Risk premia may depend on market volatility, characteristics of a particular series, and some type of benchmark portfolio. This thesis intends to shine light on time varying risk premia in metal prices with various nonlinear modeling strategies.

1.2 The Solution Strategy

Market crashes or jumps, in statistical terms excess kurtosis, can be modeled in many different ways. For example, a fat tailed distribution such as the Student t distribution, Double Exponential distribution, Gram Charlier distribution or Generalized t distribution can be used to allow for leptokurtosis. Alternatively, one can allow the variance of the distribution to change over time, as in the popular GARCH model which has been used extensively in the finance literature. As the variance varies over time, extreme observations become more probable.

The third and the most accountable solution, as will be argued in this thesis, is modeling these extreme market movements with Poisson jumps. Empirically, Poisson jumps can create a distribution as flexible as any of the fat-tailed distributions, and also allow for the possibility of non-zero skewness. Flexibility refers to the ability of the Poisson jump model to capture excess kurtosis. One of the most important benefits of using this jump model is that it allows an interpretation of leptokurtosis in terms of the frequency and size of crashes, as opposed to degrees of freedom equal to a particular number in the Student t distribution. A definite advantage is being able to predict the frequency of market crashes using the Poisson jump model.

Empirical researchers have concluded that allowing for a time-varying variance is not adequate in modeling the excess kurtosis in financial time series. Incorporation of jump components can fill this gap. Market volatility changes over time, in a manner that is dependent on its past history, as this history provides an important element for investment decisions. Similarly, market participants observe market jumps/crashes and their reaction to new unexpected information is likely to be determined by what they did in the past, which is reflected in the crashes themselves. A time dependent jump frequency may be an essential element for understanding financial series.

This thesis will also examine implications of using various types of modeling strategy to capture risk premia in metal returns. The first approach is to model risk premia in proportion to the conditional covariance between metal returns and a benchmark portfolio formulated as a bivariate GARCH-M model. The rationale is that as the market becomes more volatile, agents would demand a higher return to bear the increasing risk. Risk premia are generated by the market participants as suitable level of compensation. As an extension to the bivariate GARCH-M model, the effect of unusually large movements on risk premia is also examined by the incorporation of systematic jumps.

1.3 Contributions

The thesis is organized into three main chapters, supplemented by introductory and concluding chapters. Chapter Two develops a new conditional jump model that is used to study jump dynamics in stock market returns. We proposed a simple filter to infer *ex post* the distribution of jumps. Applying the simple Bayes rule to the Poisson distribution allows us to infer the realization of jumps after observing the returns. This permits construction of the *shock* affecting the time t conditional jump intensity (frequency), and is the main input into an autoregressive conditional jump intensity model. The model allows the conditional jump intensity to be time-varying and follows an approximate ARMA form.

The time-series characteristics of 72 years of daily stock returns are analyzed using the jump model, coupled with a GARCH specification of volatility. We find significant time-variation in the conditional jump intensity and evidence of time-variation in the jump size distribution. The conditional jump dynamics contribute to a good in-sample and out-of-sample fit to stock market volatility, and capture the rally often observed in equity markets following a significant downturn.

Chapter Three generalizes the model in a bivariate framework and develops a new bivariate Poisson jump model to study jump dynamics in foreign exchange returns. The new model extends a multivariate GARCH parameterization to include a bivariate correlated jump process. The conditional covariance matrix has the BEKK (Baba, Engle, Kraft, and Kroner 1989) structure, while the bivariate jumps are governed by a Correlated Bivariate Poisson (CBP) function. Testing for correlated jumps and time-varying jump intensity are discussed and hypothesis testing based on likelihood ratio and Ljung Box statistics are proposed.

Using daily exchange rate data for two currencies from the foreign exchange market, we find evidence of both independent currency specific jumps, as well as jumps common to both exchange rates. A relationship between the correlated jump intensity and the correlations between returns is established. The chapter concludes by investigating a time-varying structure for the arrival of jumps that relaxes the assumption of a constant and bounded jump correlation imposed by the CBP function.

Chapter Four extends the recent empirical literature on risk-adjusted Hotelling Rules for exhaustible resources by allowing time-varying risk premia in resource asset returns. Monthly metal prices for four metals (lead, copper, silver, and zinc) are used to estimate a Bivariate GARCH-M model derived from the risk-adjusted Hotelling rule and the consumption-based Intertemporal Asset Pricing Model (IAPM) of resource asset returns. The second part of the analysis augments the bivariate model with the Correlated Bivariate Poisson jump component to study the effect of extreme market movements on resource returns. Evidence of time varying risk premia is found in copper and silver prices; however, the sizes of risk premia in general are too small to be important for empirical application. The results basically reject the hypothesis that time-varying risk premia provide a key component in modeling nonrenewable resource asset returns.

In summary, this thesis proposes a set of models that are capable of

- capturing excess kurtosis often observed in financial time series, thus avoiding erroneous inferences and exploiting additional information available in the series.
- allowing the intensity of jumps to follow a learning process, capturing the possibility of market participants frequently adjusting for their mistakes.
- modeling simultaneous as well as unique jumps in multiple series, thereby gaining efficiency from this unexplored relationship.
- performing hypothesis tests on the characteristics of market crashes (such as the

presence of jumps, time varying frequency, and correlated jumps) as well as on economic theories (i.e. the Risk-Adjusted Hotelling rule in resource economics).

This research project develops a new set of econometric models with accompanying tools for interpretation, implementation, cross validation, and hypothesis testing. This research provides an avenue for further innovations in the fields of nonlinear econometrics, empirical finance, and natural resource economics.

Chapter 2

Conditional Jump Dynamics in Stock Market Returns

2.1 Introduction

Over past decades several stylized facts have emerged regarding the statistical behavior of speculative market returns. The most important of these empirical findings are: (i) asset returns are approximately a martingale difference sequence; (ii) the conditional variance is time-varying; and (iii) the unconditional distribution displays leptokurtosis. Conventional wisdom on volatility dynamics suggests that GARCH and stochastic volatility (SV) models provide a good first approximation to these stylized facts by modeling the autoregressive structure in the conditional variance. Both the GARCH and SV models are designed to capture smooth persistent changes in volatility. These models are however not suited to explaining large discrete changes which are often found in asset returns. A host of studies has investigated the nature of leptokurtosis in financial series. In most speculative markets, an allowance for discrete jumps in returns is necessary to better match statistical features observed in the data (Andersen, Benzoni, and Lund (1999) and Gallant, Hsieh, and Tauchen (1997)), as well as to reconcile mispricing in options markets (Bakshi, Cao, and Chen (1997), Bates (1996), Das and Sundaram (1999), and Jorion (1988)). A large literature has investigated the importance of jumps from a statistical and asset pricing perspective.

Allowing for jumps in asset returns introduces two additional random variables: jump intensity and jump size. The jump intensity refers to the average number of jumps that occur in the returns during a fixed time interval. The Poisson distribution is the most popular choice for modeling the intensity. For the jump size distribution, since no prior knowledge is given, the normal distribution is frequently used in the literature.

To explore the importance of time-variation in the jump intensity in a more general framework, we propose a new discrete time model in which the conditional jump intensity follows an endogenous autoregressive process. To make estimation straightforward we assume that the conditional jump intensity can be projected onto observables contained in the most recent information set. In our model, jumps are unobserved and therefore difficult to analyze directly. The first step in our approach is to propose a filter to infer *ex post* the distribution of jumps at time t. Using the filter we next construct the *shock* in the expected number of jumps as observed by the econometrician. This *shock* at time t provides the basic input into next period's conditional jump intensity. Our model of autoregressive conditional jump intensity (ARJI) specifies the conditional jump intensity follows an approximate ARMA process.

There are several advantages to our approach for exploring time-variation in the jump intensity. First, since the jump intensity has an ARMA functional form, it is possible to parsimoniously capture many forms of autocorrelation. Second, the model is easy to estimate via maximum likelihood estimation and asymptotic inference is available. A byproduct of estimation is the filter which provides *ex post* inference regarding the latent jump dynamics. Although we project the conditional jump intensity onto past observables, we expect our model will provide a good approximation to a process where the jump intensity follows a latent stochastic process. Such a model would require simulation methods for estimation, while our model does not. Finally, multi-period forecasts of future expected jumps can be directly calculated in our model.

Time-variation in the jump intensity may formulate only one part of the conditional jump dynamics in stock returns. In particular, the distribution governing the jump size may be time-varying. To investigate this we make the standard assumption that the jump distribution is normally distributed, but allow the conditional mean and variance of the jump size to be a function of observables. For example, we estimate models that permit the jump variance to be related to lagged squared returns and a GARCH variance factor. In addition, we explore whether the mean of the conditional jump distribution is asymmetrically related to the most recent positive or negative return. The motivation for this asymmetry is to consider whether jump dynamics might capture the stock market rally frequently observed after a crash.

Jump dynamics in stock market returns are studied using the ARJI model coupled with a GARCH specification, and applied to over 72 years of daily returns on the Dow Jones Industrial Average (DJIA) price index. Using the filter we find a constant jump intensity for the DJIA is violated, and a low order ARJI model adequately captures the time-variation in the conditional jump intensity.

The empirical results indicate that the autocorrelation in the conditional jump intensity in stock returns is positive and very persistent. Similar to the GARCH parameterization of volatility, a high probability of many (few) jumps today tends to be followed by a high probability of many (few) jumps tomorrow. Unconditionally jumps are infrequent; however, conditionally jumps show significant time-variation over our sample of data. For example, during the 1940s the daily jump intensity ranges from only 0.03 up to 2.02. This indicates there are periods in the 1940s where almost no jumps are expected (0.03) and periods where several jumps (2.02) are expected. Before both the 1929 (in-sample) and the 1987 (out-of-sample) stock market crashes we find evidence of an increase in the conditional expected jump intensity.

Allowing the conditional variance of the jump size distribution to be linearly related to a measure of market volatility, such as past squared returns, improves the models' in-sample fit and the out-of-sample forecasts of volatility. Our model identifies the rally after a significant stock market downturn. For instance, any decrease in the market of 2.5 percent or more, implies a positive conditional mean in next period's jump size distribution. Therefore, the day after a stock market crash, the likelihood of a jump next period does not necessarily decrease but the likelihood of a negative jump decreases, and the likelihood of a positive jump increases.

Finally, previous research (Bates (2000), and Chernov, Gallant, Ghysels, and Tauchen (1999)) has investigated whether there is a relationship between the jump intensity and a SV specification of volatility. In our model the analogous relationship is between the conditional jump intensity parameterization and the GARCH specification. In general, we found mixed results. After permitting the variance of the jump size distribution to be a function of the GARCH variance we found no evidence that the GARCH process affects the conditional intensity specification for our data set. However, in a similar specification in which lagged squared returns affect the jump size distribution rather than the GARCH variance, we found that the GARCH variance is positive and significant in affecting the conditional intensity.

Time-variation in the conditional jump intensity implies time-variation in the volatility and also in the conditional skewness and conditional kurtosis of returns. Thus, conditional jump dynamics may be important in explaining higher-order moments in speculative returns. The systematic persistence in the likelihood of jumps uncovered in this study may have important implications for forecasting volatility, risk management and derivative pricing.

This chapter is organized as follows. Section 2.2 provides a literature review. Section 2.3 presents a model of conditional jump dynamics coupled with a GARCH parameterization for financial market returns. A filter and the statistical features of the model are emphasized. Section 2.4 offers specification tests. Section 2.5 discusses competing models. Section 2.6 details the data used in the empirical application, while Section 2.7 reports the estimates and features of the conditional jump models applied to DJIA returns. Section 2.8 contains a discussion of the results. Finally Section 2.9 concludes.

2.2 Literature Review

The earliest evidence that financial asset return distributions tend to have thick tails is documented in Mandelbrot (1963) and Fama (1965). Mandelbrot (1963) shows that the empirical distribution is more "peaked" than the normal distribution using wool and copper prices. Fama (1965) (and many others) generated a large literature that modeled asset returns as i.i.d. draws from leptokurtic distributions such as the Paretian distribution.

The basic Poisson jump model of stock returns used in finance was introduced in Press (1967). Press (1967) labels his approach a compound events model, since it can be motivated as the aggregation of a random number of price changes within a fixed time interval. The Poisson distribution is assumed to govern the number of events that result in price movements, and the average number of events in a time interval is called the intensity. The model is capable of producing skewness and excess kurtosis in returns. All volatility dynamics are assumed to be the result of discrete jumps in stock returns, and the size of a jump is stochastic and normally distributed.

Several early empirical applications have shown the usefulness of the Press model. Akgiray and Booth (1988), Tucker and Pond (1988) and Hsieh (1989b) find a normal-Poisson jump model provides a good statistical characterization of daily exchange rates. Using nine years of daily data for the British pound, French franc, and German Deutsch mark, Akgiray and Booth (1988) demonstrated that these exchange rates can be modeled adequately by using a mixed jump diffusion process. Their results also show that the jump diffusion model is superior to the stable distribution and discrete mixture of normal distributions.

Tucker and Pond (1988) apply a mixed jump model to 15 years of six daily exchange rates and a thorough comparison is made with three competing models including compound normal, scaled-t, and general stable distributions. A series of pair-wise likelihood ratio tests concludes that the mixed jump model performs best in capturing the dynamics in foreign exchange rates. Hsieh (1989b) on the other hand suggests the mixed jump model as a useful alternative to the Student-t (Bollerslev (1987)), Generalized Error (GED) (Nelson (1988)) , and Normal-Lognormal mixture (Clark (1973)) distributions in the presence of GARCH effects. Similar results are found using stock returns in Ball and Torous (1983) where the presence of a Bernoulli jump process is verified with two years of daily data on 47 NYSE listed stocks.

The basic jump model has been extended in a number of directions. Estimation of continuous-time SV jump diffusion models requires simulation methods and has only recently been investigated in Andersen, Benzoni, and Lund (1999), Craine, Lochstoer, and Syrtveit (2000), Eraker, Johannes, and Polson (1999) and Chernov, Gallant, Ghysels, and Tauchen (1999). Craine, Lochstoer, and Syrtveit (2000) estimate a Stochastic-Volatility Jump-Diffusion model with Norwegian-British exchange rates using simulation based methods and find that jump components are an important feature of the data in addition to the stochastic volatility. A tractable alternative is to combine jumps with an ARCH/GARCH model in discrete time. In this case the GARCH parameterization explains the smooth changes in volatility while the jumps explain infrequent large discrete movements in the asset returns. Applications of a GARCH-jump mixture model include Jorion (1988), Vlaar and Palm (1993), and Nieuwland, Vershchoor, and Wolff (1994).

A common thread in these GARCH-jump mixture models is the assumption that a constant Poisson distribution directs the jump probability through time. However, it seems likely that the jump probability would change over time. Would we expect the probability of a jump in stock market returns prior to the 1987 stock market crash to be the same as other periods? The results in Bates (1991) would suggest the answer to this is no. Using S&P 500 futures options and assuming an underlying jump diffusion Bates (1991) finds systematic behavior in expected jumps before the 1987 crash.

Recent research has further extended the theoretical framework to permit a time-varying jump distribution. For example, Das (1998) and Fortune (1999) use dummy variables to allow the jump intensity to change over the week. Chernov, Gallant, Ghysels, and Tauchen (1999) estimate specifications which allow the jump intensity to depend on the size of previous jumps, and a stochastic volatility factor. Eraker, Johannes, and Polson (1999) model jumps in both returns and volatility.

2.3 A Dynamic Conditional Jump Model for Stock Returns

In this section we present a discrete time jump model with a time-varying conditional jump intensity and jump size distribution. First, we lay out the basic Poission Jump Model with GARCH variance. Next, we propose the new Autoregressive Jump Intensity (ARJI) model and provide a detailed discussion of properties and estimation procedures. Finally, we discuss the generalization of the jump size.

2.3.1 Basic Poisson Jump Model

Due to the vast literature showing that GARCH models provide a good first approximation to the conditional variance, we combine the jump specification with a GARCH parameterization of volatility. The GARCH model captures the stylized fact that the conditional variance is time varying and the jump component models the important leptokurtic characteristic.

Define the information set at time t to be the history of returns, $\Phi_t = \{R_t, \ldots, R_1\}$. Consider the following jump model for stock returns,

$$\begin{aligned} R_t &= \mu + \sum_{i=1}^{l} \phi_i R_{t-i} + \sqrt{h_t} z_t + \sum_{k=1}^{n_t} Y_{t,k} \\ z_t &\sim NID(0,1), \quad Y_{t,k} \sim N(\theta_t, \delta_t^2) \end{aligned} \tag{2.1}$$

where μ represents the constant in returns; ϕ_i is the autoregressive coefficient for the ith lag and i goes from 1 to l; z_t is standard normal random variable; h_t is the conditional GARCH variance which captures the time varying dynamics in conditional variance; $Y_{t,k}$ is the jump size for the kth jump at time perod t and for each time period, R_t may experience 1 to n_t jumps. The jump component $\sum_{k=1}^{n_t} Y_{t,k}$ is used to capture excess kurtosis.

The conditional jump size $Y_{t,k}$, given Φ_{t-1} is presumed to be independent, and normally distributed with mean θ_t and variance δ_t^2. To simplify construction of the likelihood we specify both z_t and the jump size $Y_{t,k}$, as independent normal random variables; however, our model of the conditional jump dynamics does not depend on this assumption.

Let n_t denote the discrete counting process governing the number of jumps that arrive between $t-1$ and t, which is distributed as a Poisson random variable with parameter $\lambda_t > 0$ and density,

$$P(n_t = j|\Phi_{t-1}) = \frac{\exp(-\lambda_t)\lambda_t^j}{j!} \quad j = 0, 1, 2, \ldots. \tag{2.2}$$

The mean and variance for the Poisson random variable are both λ_t, which is often called the (jump) intensity. We will permit the jump intensity to be time-varying. The process which λ_t follows is discussed below, but for now we assume that knowledge of the information set at time $t-1$ implies knowledge of λ_t. To complete the specification of the conditional volatility dynamics for returns, let h_t be measurable with respect to the information set Φ_{t-1} and follow a GARCH(p,q) (Bollerslev (1986)) process. That is,

$$h_t = \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i} \tag{2.3}$$

where $\epsilon_t = R_t - \mu - \sum_{i=1}^{l} \phi_i R_{t-i}$. This specification of ϵ_t contains the expected jump component and therefore allows it to propagate and affect future volatility through the GARCH variance factor.[1]

Imposing the restrictions of a constant jump intensity ($\lambda_t = \lambda$) and a constant jump size distribution ($\theta_t = \theta$, $\delta_t^2 = \delta^2$), our model nests several mixed jump models that have been investigated in the literature. For example, Jorion (1988) estimates a constant jump intensity-ARCH model for foreign exchange and stock market returns, while both Vlaar and Palm (1993) and Nieuwland, Vershchoor, and Wolff (1994) use a constant jump intensity-GARCH model to capture exchange rate dynamics.

Several extensions to this model have been proposed that permit the intensity to be time-varying, typically driven by some exogenous vector X_t thought to identify the likelihood of jumps. For instance, X_t might include dummy variables (Das (1998) and Fortune (1999)), or macro variables such as interest rates (Bekaert and Gary

[1] An alternative definition for ϵ_t that includes the conditional expectation from the jump component is $R_t - \mu - \sum_{i=1}^{l} \phi_i R_{t-i} - \lambda_t \theta_t$. In our empirical application we find the former specification to yield a substantially better log-likelihood value.

(1998) and Neely (1999)). A problem with this approach is the choice of *news* events or information to include in X_t. Rather than follow this approach we allow λ_t to endogenously evolve according to a parsimonious ARMA structure.

2.3.2 Autoregressive Jump Intensity (ARJI) Model

Consider the following model of autoregressive conditional jump intensity, denoted ARJI(r,s). Let $\lambda_t \equiv E[n_t|\Phi_{t-1}]$ be the conditional expectation of the counting process which is assumed to follow,

$$\lambda_t = \lambda_0 + \sum_{i=1}^{r} \rho_i \lambda_{t-i} + \sum_{i=1}^{s} \gamma_i \xi_{t-i}. \tag{2.4}$$

The conditional jump intensity at time t is related to r past lags of the conditional jump intensity plus lags of ξ_t. ξ_{t-i} represents the innovation to λ_{t-i} as measured *ex post* by the econometrician. This *shock*, or jump intensity residual, is calculated as follows,

$$\begin{aligned} \xi_{t-i} &\equiv E[n_{t-i}|\Phi_{t-i}] - \lambda_{t-i} \\ &= \sum_{j=0}^{\infty} jP(n_{t-i} = j|\Phi_{t-i}) - \lambda_{t-i} \end{aligned} \tag{2.5}$$

The first term on the right hand side of Equation 2.5 is our inference on the average number of jumps at time $t-i$ based on time $t-i$ information, while the second term in (2.5) is our expectation of the number of jumps using information at time $t-i-1$. Therefore, ξ_{t-i} represents the unpredictable component affecting our inference about the conditional mean of the counting process n_{t-i}.

Let $f(R_t|n_t = j, \Phi_{t-1})$ denote the conditional density of returns given that j jumps occur given the information set Φ_{t-1}. Having observed R_t and using Bayes' rule we can infer *ex post* the probability that a jump of size j occurred at time t, with the filter defined as,

$$P(n_t = j|\Phi_t) = \frac{f(R_t|n_t = j, \Phi_{t-1})P(n_t = j|\Phi_{t-1})}{P(R_t|\Phi_{t-1})}, \quad j = 0, 1, 2, \ldots \tag{2.6}$$

where $P(n_t = j|\Phi_{t-1})$ is given by Equation 2.2. The filter in Equation 2.6 is an important component of our model of time-varying jump dynamics because of its role in Equation 2.5 and also because it can be constructed for purposes of inference. For example, the probability that at least one jump occurred can be assessed as $1 - P(n_t = 0|\Phi_t)$. The filter may be particularly useful in revealing misspecification in the simpler constant intensity specification. The constant intensity specification assumes that the probability of jump is constant over time which implies that the difference between the probabilities produced by the filter and the Poisson density should be zero.

The conditional density of returns is completed by integrating out the discrete valued variable n_t, governing the number of jumps. This conditional density of returns can be written as,

$$P(R_t|\Phi_{t-1}) = \sum_{j=0}^{\infty} f(R_t|n_t = j, \Phi_{t-1}) P(n_t = j|\Phi_{t-1}). \tag{2.7}$$

Equation 2.7 shows that this model is nothing more than a discrete mixture of distributions, where the mixing is driven by a time-varying Poisson distribution. The assumptions in Equation 2.1 imply that the distribution of returns conditional on the most recent information set and j jumps is normally distributed as,

$$f(R_t|n_t = j, \Phi_{t-1}) = \frac{1}{\sqrt{2\pi(h_t + j\delta_t^2)}} \exp\left(-\frac{(R_t - \mu - \sum_{i=1}^{l} \phi_i R_{t-i} - \theta_t j)^2}{2(h_t + j\delta_t^2)}\right). \tag{2.8}$$

Construction of the likelihood and maximum likelihood estimation follow by iterating on Equations 2.4, 2.6 and 2.7. Equation 2.7 involves an infinite sum over the possible number of jumps n_t. In practice we truncate the maximum number of jumps to a large value τ, so that the probability of τ or more jumps is zero. This is empirically checked for a particular dataset and set of estimates in that, to machine precision, Equation 2.2 is 0 for $j \geq \tau$. A second check on the choice of τ is to consider $\tilde{\tau} > \tau$ to ensure that the likelihood and parameter estimates do not change.

In this model the conditional jump intensity is time-varying and, under certain circumstances enunciated below will have an unconditional value. To derive the unconditional value of λ_t, first note that ξ_t is a martingale difference sequence with respect to Φ_{t-1}, since,

$$E[\xi_t|\Phi_{t-1}] = E[E[n_t|\Phi_t]|\Phi_{t-1}] - \lambda_t = \lambda_t - \lambda_t = 0 \tag{2.9}$$

and therefore, $E[\xi_t] = 0$, and $\mathrm{Cov}(\xi_t, \xi_{t-i}) = 0, i > 0$. Another way to see this result is to note that ξ_t is by definition nothing more than the rational forecast error associated with updating the information set. This is, $\xi_t = E[n_t|\Phi_t] - E[n_t|\Phi_{t-1}]$.

By using the approximate ARMA form for the evolution of λ_t from Equation 2.4 many of the results for ARMA models are directly applicable to this model. For example, assuming the roots of the polynomial $(1 - \sum_{i=1}^{r} \rho_i L^i)$, where L is the lag operator associated with Equation 2.4, lie outside the unit circle then the unconditional value of λ_t exists and is,

$$E[\lambda_t] = \frac{\lambda_0}{1 - \sum_{i=1}^{r} \rho_i}. \tag{2.10}$$

Furthermore, conditional forecasts of the future jump intensity can be formed using Equation 2.4. To illustrate, consider the case where $r = s = 1$. Then

$$E[\lambda_{t+i}|\Phi_{t-1}] = \begin{cases} \lambda_t & i = 0 \\ \lambda_0(1 + \rho + \ldots + \rho^{i-1}) + \rho^i \lambda_t & i \geq 1. \end{cases} \tag{2.11}$$

Recall that λ_t is measurable with respect to the information set Φ_{t-1}. If $|\rho| < 1$, then as i becomes large the forecast approaches the unconditional value in (2.10).

For the Poisson distribution to be well defined λ_t must be positive. Note that in the case of $r = s$, Equation 2.4 can be rewritten as

$$\lambda_t = \lambda_0 + \sum_{i=1}^{r} (\rho_i - \gamma_i)\lambda_{t-i} + \sum_{i=1}^{r} \gamma_i E[n_{t-i}|\Phi_{t-i}]. \tag{2.12}$$

Subject to reasonable startup conditions that ensure $\lambda_i > 0,\ \ i = -r+1, \ldots, 0$, then a sufficient condition for $\lambda_t > 0$ for all t is $\lambda_0 > 0,\ \rho_i > \gamma_i$, and $\gamma_i > 0$. To estimate the ARJI model, startup values of λ_i and ξ_i for $i \leq 0$ must be set. In our empirical application discussed in the next section, we set startup values of the jump intensity to the unconditional value in Equation 2.10 and $\xi_i = 0$. Alternatively these values could be estimated or the values could be arbitrarily set, since asymptotically the first observation is negligible to the likelihood function.

Consider the intuition behind the evolution of the conditional jump intensity in the ARJI model. Suppose we observe $\xi_t > 0$ for several periods. This suggests that the jump intensity is temporarily trending away from its unconditional mean. This model effectively captures systematic changes in jump risk in the market. The likelihood of large discrete changes in foreign exchange markets or crashes in stock markets may change considerably over time. The ARJI model can capture systematic changes and also forecast increases (decreases) in jump risk into the future.

The ARJI model of λ_t is convenient from both an estimation and forecasting perspective and should be useful in capturing time-series dynamics similar to those which an ARMA model can capture. As we show in the empirical section the linear specification for λ_t appears to work well for daily stock returns. However, other functional forms that include nonlinearity may be very useful. In this case lags of λ_t, ξ_t, and other variables in the information set may enter a nonlinear function driving the conditional intensity parameter at time t.

The time-series model of λ_t is not a true ARMA model in that it is not driven by an unforecastable innovation, but rather a measurable one with respect to Φ_{t-1}. However, we would expect that this model will provide a good approximation if λ_t did follow a true latent ARMA structure. Such a model would require simulation methods to compute the likelihood while ours does not.

2.3.3 Time Varying Jump Size

Up until this point our model has focused on the conditional dynamics governing the number of jumps. However, the distribution of the jump size which is postulated to be Gaussian may also change and display conditional dynamics. To explore this further we consider the following two extensions of the model. The first allows the conditional mean and conditional variance of the jump size distribution to be conditionally normal and a function of past returns,

$$\theta_t = \eta_0 + \eta_1 R_{t-1} D(R_{t-1}) + \eta_2 R_{t-1}(1 - D(R_{t-1})) \tag{2.13}$$

$$\delta_t^2 = \zeta_0^2 + \zeta_1 R_{t-1}^2 \tag{2.14}$$

where $D(x) = 1$, $x > 0$ and otherwise 0, and $\eta_0, \eta_1, \eta_2, \zeta_0$, and ζ_1 are parameters to be estimated. This specification of the conditional mean of the jump size allows some flexibility concerning where jumps are centered. For example, if during the last period the market had experienced a gain (decline) then today's conditional mean of the jump size would be $\eta_0 + \eta_1 R_{t-1}$, $(\eta_0 + \eta_2 R_{t-1},)$. Thus, the first moment of the jump distribution can respond to whether last period's market return was positive or negative and to its magnitude. This formulation may explain the rally after a stock market crash through a change in the jump direction. For this, we would expect $\eta_2 < 0$. To investigate whether the jump size variance might be sensitive to the overall level of market volatility we allow R_{t-1}^2 to affect δ_t^2. We label the extension in (2.13)-(2.14) as ARJI-R_{t-1}^2.

A second extension of interest is whether the variance of the jump size is a function of the GARCH volatility. The formulation for θ_t is the same as in (2.13), but now,

$$\delta_t^2 = \zeta_0^2 + \zeta_1 h_t, \tag{2.15}$$

which we refer to as ARJI-h_t. The difference between these two specifications (2.14 and 2.15) of the jump size variance is that while the lagged squared return is a proxy for last period's market volatility, h_t is a prediction of the time t GARCH volatility component of our model. If the variance of the jump size is sensitive to contemporaneous market volatility then this final specification of the conditional variance of jumps may capture this effect better than (2.14).

To derive the conditional mean and variance of our model first redefine (2.1) so that $R_t = B_t + C_t$ where $C_t = \sum_{k=1}^{n_t} Y_{t,k}$, is the jump component and $B_t = R_t - C_t$ is the remainder. Since our GARCH-jump model is a discrete mixture of distributions the ith uncentered moment of C_t is,

$$E[C_t^i|\Phi_{t-1}] = \sum_{j=0}^{\infty} E[C_t^i|n_t = j, \Phi_{t-1}]P(n_t = j|\Phi_{t-1}), \quad i > 0. \tag{2.16}$$

Then the first two conditional uncentered moments (See Appendix) of C_t are,

$$E[C_t|\Phi_{t-1}] = \sum_{j=0}^{\infty} j\theta_t P(n_t = j|\Phi_{t-1}) = \theta_t\lambda_t \tag{2.17}$$

and

$$\begin{aligned} E[C_t^2|\Phi_{t-1}] &= \sum_{j=0}^{\infty}(j(\delta_t^2 + \theta_t^2) + \theta_t^2(j^2 - j))P(n_t = j|\Phi_{t-1}) \\ &= (\delta_t^2 + \theta_t^2)\sum_{j=0}^{\infty} jP(n_t = j|\Phi_{t-1}) + \theta_t^2\sum_{j=0}^{\infty}(j^2 - j)P(n_t = j|\Phi_{t-1}) \\ &= (\delta_t^2 + \theta_t^2)\lambda_t + \theta_t^2\lambda_t^2 \end{aligned}$$

and hence

$$\mathrm{Var}(C_t|\Phi_{t-1}) = (\delta_t^2 + \theta_t^2)\lambda_t. \tag{2.18}$$

Using these results the conditional mean and variance of the returns are

$$\begin{aligned} E[R_t|\Phi_{t-1}] &= \mu + \sum_{i=1}^{l}\phi_i R_{t-i} + \theta_t\lambda_t & (2.19) \\ \mathrm{Var}(R_t|\Phi_{t-1}) &= h_t + (\delta_t^2 + \theta_t^2)\lambda_t. & (2.20) \end{aligned}$$

Note that time-variation in λ_t and the conditional jump size distribution affect both the conditional mean and conditional variance of the returns. The conditional variance of returns is an increasing function in the jump intensity while the conditional mean of returns can be increasing or decreasing depending on the sign of θ_t. The conditional jump dynamics also imply conditional skewness and kurtosis. Similar calculations (See Appendix A), or drawing on the results in Das and Sundaram (1997) show the conditional skewness (Skew) and conditional kurtosis (Kurt) to be,

$$\text{Skew}(R_t|\Phi_{t-1}) = \frac{\lambda_t(\theta_t^3 + 3\theta_t\delta_t^2)}{(h_t + \lambda_t\delta_t^2 + \lambda_t\theta_t^2)^{3/2}} \tag{2.21}$$

$$\text{Kurt}(R_t|\Phi_{t-1}) = 3 + \frac{\lambda_t(\theta_t^4 + 6\theta_t^2\delta_t^2 + 3\delta_t^4)}{(h_t + \lambda_t\delta_t^2 + \lambda_t\theta_t^2)^2}. \tag{2.22}$$

2.4 Specification Tests

In order to evaluate the in-sample and out-of-sample performance of the proposed models we consider several specification tests that are based on the integral transformation of Rosenblatt (1952).

Consider a candidate model with conditional density $\tilde{f}(R_t|\Phi_{t-1}, \Theta)$, where Θ is the known parameter vector. Rosenblatt (1952) shows that if the data are drawn from the density $\tilde{f}(\cdot)$ then the series $\{\tilde{u}_t\}_{t=1}^T$, defined from,

$$\tilde{u}_t = \int_{-\infty}^{R_t} \tilde{f}(v|\Phi_{t-1}, \Theta)dv \tag{2.23}$$

will be uniform *i.i.d.* or $U(0,1)$. Therefore testing whether or not $\tilde{u}_t \sim i.i.d.\ U(0,1)$ provides a test for correct model specification.

The first test we consider is a Pearson goodness-of-fit test for $\tilde{u}_t$ suggested by Vlaar and Palm (1993). Under the null hypothesis of a correctly specified model,

$$\sum_{i=1}^{g} \frac{(n_i - En_i)^2}{En_i} \sim \chi^2(g-1), \tag{2.24}$$

where g is the number of equally spaced groups, and n_i is the number of observation of $\tilde{u}_t$ that occur in group i. After g is chosen, $n_i = \sum_{t=1}^T I_{it}$ where,

$$I_{it} = \begin{cases} 1 & \text{if } (i-1)/g < u_t \leq i/g \\ 0 & \text{otherwise} \end{cases} \tag{2.25}$$

for $1 \leq i \leq g$.

To focus on the dynamics of the conditional distribution, Palm and Vlaar (1997) and Berkowitz (2001) recommend applying an inverse normal transformation to the data $\{\tilde{u}_t\}_{t=1}^T$ in order to test if the transformed data $\{\tilde{v}_t\}_{t=1}^T$ are independent standard normal. Under the null hypothesis of a correctly specified model, $\tilde{v}_t$ is normally identically distributed or $NID(0,1)$. Let $\tilde{v}_t = F^{-1}(\tilde{u}_t)$, where $F^{-1}(\cdot)$ is the inverse of the standard normal CDF. We consider a likelihood ratio (LR) test for $a_0 = a_1 = \cdots = a_5 = 0, \sigma = 1$, in the following regression,

$$\tilde{v}_t = a_0 + a_1\tilde{v}_{t-1} + \cdots + a_5\tilde{v}_{t-5} + \sigma\nu_t \tag{2.26}$$

against an unrestricted alternative hypothesis that only maintains that ν_t is normally distributed. Under the null hypothesis of a correctly specified model the LR statistic is distributed as $\chi^2(7)$.

The final test is based on Diebold, Gunther, and Tay (1998) and evaluates graphically whether or not $\tilde{u}_t \sim i.i.d.\ U(0,1)$. This is done through a histogram of $\{\tilde{u}_t\}_{t=1}^T$ and estimates of the autocorrelation functions of successive powers of $\tilde{u}_t$. For a correctly specified model $\tilde{u}_t$ will have the density of the uniform distribution while the powers of $\tilde{u}_t$ should display no evidence of autocorrelation. Like the Pearson goodness-of-fit test, the density estimate of $\tilde{u}_t$ provides a measure of a model's ability to capture the unconditional distribution of returns, whereas the estimates of the autocorrelation functions of powers of $\tilde{u}_t$ assess the adequacy of the conditional distribution. It should be noted that all of the tests in this section ignore parameter uncertainty and in practice the tests are applied to models with an estimated parameter vector.

2.5 Alternative Models

The ARJI model is just one of the many alternatives available to model excess kurtosis. In theory, any thick tail distribution can be used to capture leptokurtosis. For example, The Student-t, Gram Charlier, Power Exponential, and Generalized Student-t distributions are very popular in the literature. We will estimate these models and compare the results of out-of-sample volatility forecasts to the ARJI model in Section 2.7. This section will briefly discuss the properties of each competing model.

2.5.1 Student-t Distribution

The Student-t distribution is one of the most popular distributions for dealing with excess kurtosis. A GARCH model with Student-t distribution has proven to be superior to the normal distribution in many applications (Bollerslev 1986, Bauer, Nieuwland, and Verschoor 1994). The Student t density function is defined as

$$\begin{aligned} f_{St}(u_t) &= \Gamma\left(\frac{v+1}{2}\right)\cdot\Gamma\left(\frac{v}{2}\right)^{-1}\cdot((v-2)h_t^2)^{-1/2}\times \\ & \quad (1+u_t^2h_t^{-2}(v-2)^{-1})^{(v+1)/2} \end{aligned} \tag{2.27}$$

where v represents the degrees of freedom and Γ is the gamma function. This distribution allows the error term to have a thick tail probability governed by the degrees of freedom parameter. The smaller the value of v, the higher is the probability of obtaining large realizations of u_t. As $1/v$ approach 0, this distribution will be identical to the normal distribution. Since the Student t distribution is symmetric, the possibility of skewness is eliminated.

2.5.2 Gram Charlier Distribution

To allow for both skewness and leptokurtosis, a GARCH model combined with the Gram Charlier distribution has been proposed by Lee and Tse (1991). The Gram Charlier density is defined as

$$GC(u_t) = \phi(u_t)\varphi(u_t) \tag{2.28}$$

$$\varphi(u_t) = (1+\lambda_3H_3(u_t)/6+\lambda_4H_4(u_t)/24) \tag{2.29}$$

$$H_3(u_t) = u_t^3 - 3u_t \tag{2.30}$$

$$H_4(u_t) = u_t^4 - 6u_t^3 + 3. \tag{2.31}$$

where $\phi(u_t)$ is the standard normal density and H is a Hermite polynomial. The Gram Charlier denisty has a straight forward interpretation. For example, λ_3 is the standardized measure of skewness and λ_4 represents the excess kurtosis. The Gram Charlier density collapses to the normal density when both λ_3 and λ_4 are equal to zero. This distribution can also be generalized to incorporate the higher moments. Details can be found in Kendall and Stuart (1969).

The limited applications of this distribution in the literature may be attributable to difficulties associated with estimation. Note that the GC function is well defined if and only if GC integrates to one and remains positive at all times. Problems can arise with respect to the second component, $\varphi(u_t)$ depending on the parameters (λ_3 and λ_4) as well as the Hermite polynomials (H_3 and H_4). Imposing restrictions on the parameters provide a possible solution. However, this may implicitly impose unreasonable assumptions on the skewness coefficient. Alternatively, one may use the semi-non-parametric (SNP) approach developed by Gallant and Tauchen (1989) to avoid this problem. For the purpose of this chapter, we estimate this Gram Charlier distribution using maximum likelihood methods with no restrictions imposed in order to preserve the interpretation of λ_3 and λ_4. The same strategy is adopted in Lee and Tse (1991).

2.5.3 Exponential Power Distribution

The Exponential Power distribution was first introduced by Box and Tiao (1973). The density function is given by

$$EP(u_t) = w(u_t)exp[-c(\gamma)|u_t|^{1/\gamma}] \tag{2.32}$$

$$w(\gamma) = \frac{1}{2\gamma}\left[\frac{\Gamma(3\gamma)}{\Gamma^3(\gamma)}\right]^{1/2} \quad (2.33)$$

$$c(\gamma) = \left[\frac{\Gamma(3\gamma)}{\Gamma(\gamma)}\right]^{\frac{1}{2\gamma}}. \quad (2.34)$$

The main advantage of this distribution is the fact that it nests several distributions as special cases. For example, the EP distribution becomes the normal density when $1/\gamma$=2.0; the Double Exponential (Laplace) distribution when $1/\gamma = 1$; and the uniform distribution when $1/\gamma = \infty$. Estimation of this model by maximum likeihood estimation is straightforward, and the only restriction imposed is $\gamma > 0$.

2.5.4 Generalized t Distribution

McDonald and Newey (1988)'s Generalized t (GT) distribution goes one step further by nesting the Exponential Power, Normal, Cauchy, Double Exponential (Laplace), and Student t distribution as special and limiting cases. The density function is defined as

$$GT(u_t) = \frac{p}{2h_t q^{1/p} B(1/p, q)\left(1 + \frac{|u_t|^p}{q h_t^p}\right)^{q+\frac{1}{p}}} \quad (2.35)$$

where B is the Beta function, and p and q are parameters that control the shape of the density. The larger the values of p and q, the smaller is the probability of the occurrence of extreme values (thinner tails). The kth moments will only exist when k<pq.

The Box and Tiao (1973) Exponential Power distribution is a limiting case of the Generalized t when q approaches infinity. When $q \to \infty$ and p=1, the GT distribution is identical to the Double Exponential. The normal distribution with variance α^2 can be obtained by setting $q \to \infty$, and letting p=2, and $h_t = \sqrt{2}\alpha$. The Student t distribution can be derived from the normal distribution without $q \to \infty$. The Cauchy distribution can be derived from the Student t by setting $q = 1/2$.

2.6 Data

Table 2.1: Summary Statistics (1928 - 1984)

	Mean	Standard Deviation	Skewness	Excess Kurtosis	Q(15)
R_t	0.0107	1.1167	0.0565	18.5198	148.00
	(0.0090)	(0.00189)	(0.2923)	(2.4160)	[0.00]
$\|R_t\|$	0.7130	0.8595	4.1255	34.4842	20400.00
	(0.0069)	(0.0202)	(0.2893)	(5.0985)	[0.00]
R_t^2	0.0124	0.0522	18.3919	527.512	9620.00
	(0.0004)	(0.0048)	(1.7728)	(79.8848)	[0.00]

Summary statistics for daily Dow Jones Industrial Average (DJIA) returns from Oct 1, 1928 to Dec 31, 1984. Q(15) are modified Ljung-Box Statistics robust to heteroskedasticity for serial correlation with 15 lags. Standard errors robust to heteroskedasticity are in parentheses, and p-values are in square brackets.

The data consist of daily close, intraday high and low for the Dow Jones Industrial Average (DJIA). Returns are defined as 100 times the first-difference in the logarithm of the close of the DJIA index from October 1, 1928 to January 11, 2000. The data set contains 18,947 observations in total.

Summary statistics are reported in Table 2.1 for the sample from 1928 until the end of 1984. Standard errors that are robust to the heteroskedasticity are provided in parentheses. Although the skewness coefficient is not significantly different from zero, the deviation from normality is apparent in the excess kurtosis which is 18.52. Evidence of time dependence is found using the modified Ljung-Box statistic (West and Cho (1995)) which is robust to heteroskedasticity and reported for autocorrelations up to 15 lags in the last column of Table 2.3. The modified Ljung-Box (LB) statistics show strong serial correlations in both the levels and the squares of the return series. This is consistent with the results in Brock, Lakonishok, and LeBaron (1992) which show that the serial correlations in DJIA returns is significant but unstable and depends on the sample period.

The statistical properties of the data series can also be examined by the visual plot in Figure 2.3. The dramatic movements in the 1930s are clearly shown. Volatility clustering is unquestionable as high volatility this period tends to be followed by high volatility next period. The GARCH variance structure is appropriate for modeling this phenomenon.

Figure 2.1: Density Plot

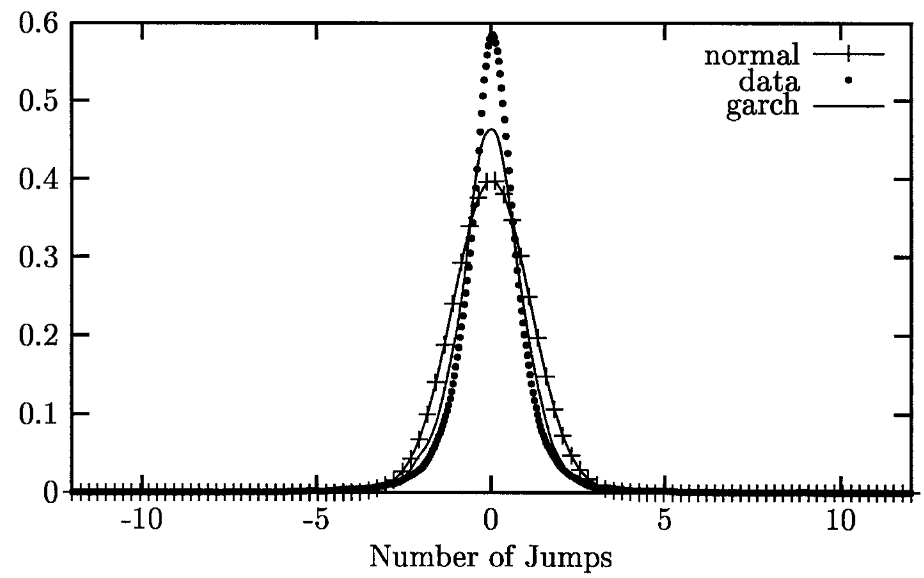


Figure 2.1 provides a denisty plot for the data, the normal distribution, and the GARCH model. The curve with the highest peak is derived from the dataset. There is no surprise that the data does not match the normal distribution (the lowest curve) well as, excess kurtosis is reported in the summary statistics. A similar comparsion is performed with the GARCH residuals. A GARCH(1,1) model is estimated and the density is plotted on the same diagram and note that it shows significant improvement compared to the data. However, there is still considerable distance between the two curves implying room for improvement.

2.7 An Application to Stock Market Returns

2.7.1 Estimation Results

Estimation of all models is conducted with data from 1928 until the end of 1984, and the remaining data from 1985-2000 is reserved for out-of-sample analysis. For all models a $\tau = 20$ was selected as the truncation point for the distribution determining the number of jumps, based on the selection method discussed in Section 2.3. Table 2.5 presents the estimation results for the GARCH-constant jump intensity model over various sub-samples while Table 2.6 presents the results for all models over the full in-sample period. Misspecification tests based on the modified Ljung Box (LB) statistic are reported for autocorrelation in the squared standardized residuals (Q^2) and the jump intensity residuals (Q_{ξ_t}) for 15 lags at the bottom of each table. We found it necessary to use an AR(2) to capture autocorrelation in the conditional mean of stock returns for all models. Similarly, all models use a GARCH(1,1) and where appropriate an ARJI(1,1) specification which LB statistics suggest is adequate for the in-sample period. In the following we use ARJI to denote the ARJI(1,1) model.

Estimates for the constant jump intensity model over the different sample periods, 1928-50, 1951-69, and 1970-84 are displayed in Table 2.5. This model imposes the restrictions $\lambda_t = \lambda, \theta_t = \theta$, and $\delta_t^2 = \delta^2$ in the ARJI model. These results provide evidence of changing jump dynamics over time. For example, the jump intensity parameter λ is 0.1512 for the full sample (reported in Table 2.6 under constant) but varies substantially across different sub-samples. λ is 0.1116 in 1928-50, and increases to 1.6742 in 1951-69. In the 1951-69 period the estimates indicate that the jump component is more important while the GARCH effects diminish compared to other period results. Both θ and δ in Table 2.5 display instability over the sub-samples. δ is estimated as high as 1.4624 and as low as 0.2711.

Evidence of time-variation in λ is also supported by the Ljung-Box statistics for

ξ_t. Recall that ξ_t is the measurable shock constructed by the econometrician using the *ex post* filter. In a correctly specified model ξ_t should not display any systematic behavior, otherwise it could be exploited to improve the model. We test for dependence in ξ_t in the constant jump model using the modified LB statistics denoted Q_{ξ_t} in Tables 2.5 and 2.6. In Table 2.6 the Q_{ξ_t} statistic rejects the constant intensity assumption.

Table 2.6 reports a series of model estimates for the simplest constant intensity jump model, the ARJI model with a constant jump size distribution, and the fully dynamic jump models, ARJI-h_t and ARJI-R^2_{t-1}. The log-likelihood for the constant intensity jump model is -18315.61 which represents an increase of 356.0 compared to a plain AR(2)-GARCH(1,1) model (estimates not reported) with no jumps. This suggests jumps in the DJIA may be important. Moreover, the ARJI parameterization, which allows the conditional jump intensity to vary over time provides an improvement in the likelihood as compared to the constant intensity model. The likelihood ratio (LR) test of a constant jump intensity ($\gamma = 0$) against the ARJI specification is 79.02. This test is nonstandard since under the null hypothesis the constant jump intensity is unidentified (under the null hypothesis if $|\rho| < 1$, then $\lambda_t = \lambda_0/(1-\rho)$).[2] Although methods such as Davies (1987) and Hansen (1996) could be employed to obtain a p-value, the magnitude of the test statistic suggests that the ARJI parameterization provides a significant statistical improvement over the constant jump

[2]Given that the time varying intensity is defined by

$$\lambda_t = \lambda_0 + \rho\lambda_{t-1} + \gamma\xi_{t-1}$$

which can be rewritten as

$$\lambda_t = \frac{\lambda_0}{1-\rho} + \gamma[\xi_{t-1} + \rho\xi_{t-2} + \rho^2\xi_{t-3} + ...]$$

Therefore, testing the presence of time varying jump intensity can be formulated as testing H_0: $\gamma = 0$

intensity model. In addition, the ARJI model captures the autocorrelation in ξ_t that was found in the constant intensity model. The p-value for the Q_{ξ_t} statistic is 0.01 for the constant intensity model while it is 0.83 for the ARJI model.

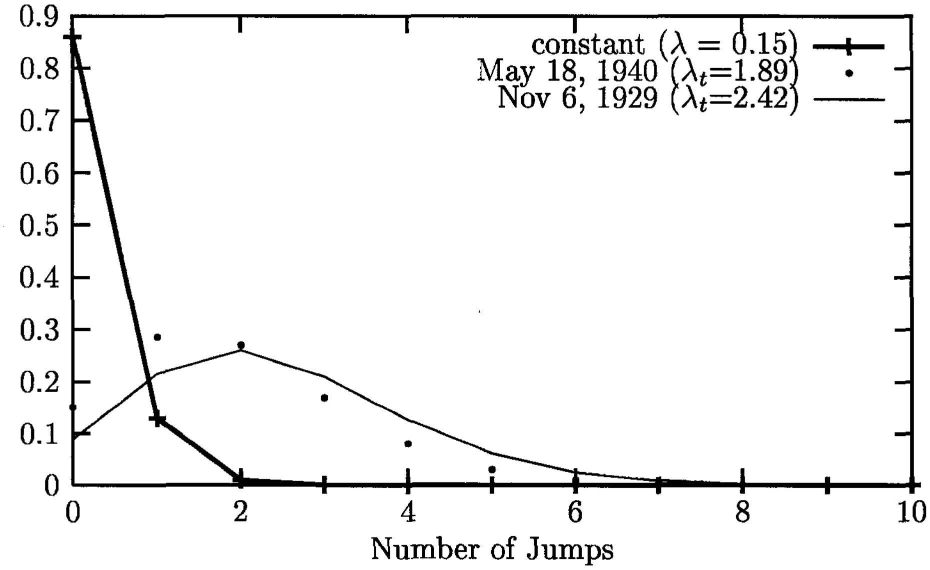
The ρ parameter in the ARJI model is estimated to be .9153 with an asymptotic standard error of 0.02 and provides a measure of the persistence in the conditional jump intensity. This suggests that a high probability of many (few) jumps today tends to be followed by a high probability of many (few) jumps tomorrow. However, unconditionally, jumps are infrequent. The unconditional jump intensity (defined in (2.10)) is only 0.1547, which is very close to the λ reported in the constant intensity model in Table 2.6. γ measures the sensitivity of λ_t to the past *shock*, ξ_{t-1}. A unit increase in ξ_{t-1} results in a dampened effect of only 0.5 on next period's jump intensity.

Figure 2.3 (see end of chapter) displays returns, the conditional intensity, and the conditional standard deviation for the ARJI specification. This figure shows considerable variation in the conditional jump intensity. To illustrate, consider the 1940s where the daily jump intensity ranges from only 0.03 up to 2.12. This indicates that there are periods in the 1940s where almost no jumps are expected (0.03) and periods where several jumps (2.12) are expected.

What effect does time-variation in λ_t have on the distribution of the number of jumps? Figure 2.2 provides two snapshots of the Poisson distribution governed by λ_t for two dates in our sample period against the Poisson distribution with the constant intensity assumption $\lambda_t = \lambda$. This figure shows that small changes in λ_t can have important effects on the Poisson distribution. Furthermore, the risk associated with the realization of jumps in the constant intensity model is considerably understated compared to the ARJI model as depicted in this figure.

Figure 2.4 displays the predictive content that λ_t has in forecasting a jump around

Figure 2.2: $P(n_t = j|\Phi_{t-1})$ from the ARJI and constant intensity model



the 1929 crash. In-sample, this model suggests that days before the crash on October 28th at least one jump was expected. A 95% confidence interval for the effect of one jump on returns is $(-2.7, 1.91)$. In fact, two or three jumps were not unlikely, since their probability of occurring just days (October 24th) before the crash was .22 and .08 respectively.

The last two columns of Table 2.6 report estimates for two models that extend the jump dynamics in the ARJI specification. Both the ARJI-R_{t-1}^2 and ARJI-h_t models allow the conditional mean and conditional variance of the jump size distribution to be a function of past returns. As measured by LR tests both models provide a significant improvement over the simpler ARJI specification. Also note these extensions do not appear to alter the dynamics found in the conditional intensity. For example, λ_t is still very persistent and ρ is .91 and .84 in the ARJI-R_{t-1}^2 and ARJI-h_t formulations, respectively.

Our estimates of these final models provide evidence that the jump direction is asymmetric and sensitive to the state of the stock market. In both the ARJI-R_{t-1}^2

Table 2.2: Specification tests

	constant	ARJI	ARJI-h_t	ARJI-R^2_{t-1}
In-Sample				
Pearson goodness-of-fit	163.930	171.945	126.798	142.564
	[4.45e-5]	[7.74e-6]	[0.031]	[0.002]
LR test	18.798	10.380	11.428	12.167
	[0.008]	[0.168]	[0.121]	[0.095]
Out-of-Sample				
Pearson goodness-of-fit	117.982	105.338	82.927	114.185
	[0.0002]	[0.003]	[0.121]	[0.0005]

P-values appear in square brackets. The Pearson goodness-of-fit test statistics are based on the integral transformation of the observed data using the respective model. Under the null hypothesis of a correctly specified model the test statistic is distributed as $\chi^2(g-1)$, where g is the number of bins. For the in-sample period (15149 observations) g=100, for the the out-of-sample period (3798 observation) g=70. The LR test is a likelihood ratio test applied to $\tilde{v}_t$ which should be distributed as $NID(0,1)$ if the conditional distribution is correctly specified. The alternative is an unrestricted AR(5) model. See Section 2.4 for details.

and ARJI-h_t specifications, θ_2 is significantly negative which implies that after a stock market downturn, the direction of a jump next period is more likely to be positive than negative.

The contribution from jumps to the conditional mean and conditional variance (Equations 2.17 and 2.18) of returns is shown in Figure 2.5 for the ARJI-h_t specification. Generally, the effect on the first two moments is small, but occasionally can be large, particularly for the conditional variance. Note that the conditional mean is affected by the asymmetric jump direction.

Given the large data set used in this study it is not surprising that the Pearson goodness-of-fit test statistics in Table 2.2 identify problems in all models, although the p-value for the ARJI-h_t formulation is only 0.031. However, the addition of jumps to a plain GARCH(1,1) model does result in a dramatic improvement in the test values (a GARCH(1,1) model with normal innovations has a p-value 6.3e-05). The jump models may not be doing a good job in fitting the unconditional distribution of returns, but the LR test in Table 2.2 indicates that the addition of the jump dynamics improves the specification of the conditional distribution in all models as compared to the constant intensity model.

2.7.2 Out-of-Sample Analysis

This section discusses and evaluates the models in the out-of-sample period, 1985-2000. In all cases, model parameters were initially set at the values in Table 2.6, thereafter every 250 observations, the model parameters are updated by estimation which included the most recently available data.

The last row of Table 2.2 contains Pearson goodness-of-fit test statistics for the out-of-sample predictive density of the models. Similar to the in-sample results the ARJI-h_t model performs the best among the group with the ARJI model ranking second with a p-value of 0.003. Figures 2.10 through 2.12 display the results from the Diebold, Gunther, and Tay (1998) evaluation of the forecast density for a GARCH(1,1) model with normal innovations, ARJI and ARJI-h_t models, respectively. Recall that, under a correctly specified model, the u_t derived from the integral transformation should be *i.i.d.* $U(0,1)$ and display no autocorrelation. Panel (a) of these figures presents an estimate of the density of u_t. The benchmark GARCH(1,1) model shows obvious problems in matching the unconditional features of the data. Both the ARJI and ARJI-h_t models provide successive improvements compared to the GARCH density. Estimates of the autocorrelations of powers of u_t are found in panels (b)-(e) of each figure. All models perform well in capturing the dynamics in the conditional forecast density, although there is some marginal improvement from adding the jump dynamics to the GARCH model.

Out-of-sample one-step ahead forecasts of volatility are evaluated against the Parkinson (1980) range statistic. This is the intraday logarithm of the ratio of the

Table 2.3: Out-of-sample forecasting performance of conditional volatility, R^2

Sample period	GARCH(1,1)	constant	ARJI	ARJI-h_t	ARJI-R_{t-1}^2
1985-2000	0.269	0.262	0.264	0.307	0.392
1990-2000	0.334	0.334	0.334	0.359	0.364

This table shows the R^2 from a regression of the range statistic on a constant and the out-of-sample, one-period ahead forecast of the conditional standard deviation for the respective models. The second column contains results for a GARCH model estimated assuming normal innovations.

high to the low transaction price and is a measure of the daily latent volatility. Traditional measures of volatility such as squared returns are very noisy and therefore practically useless as a measure of forecasting performance. The range has been used in many studies and the simulation results in Andersen and Bollerslev (1998) (see their footnote 20) show it to be a very efficient estimator compared to daily squared returns.

A good model should not only explain the variation of the data series within sample, but also accurately forecast the series out-of-sample. Figure 2.6 depicts the out-of-sample one-step ahead forecasts of λ_t for the ARJI-h_t model. Table 2.3 shows the R^2 from a linear regression of the range on a constant and the one-period ahead predicted standard deviation from the models. The first row of Table 2.3 is over the period 1985-2000 and the second row is over the period 1990-2000. Only in the second sample period does the ARJI provide a marginal improvement in out-of-sample forecasting performance as measured by R^2. However, both the ARJI-R_{t-1}^2 and ARJI-h_t models perform much better than the models without jump size dynamics. Since the jump models are designed to explain extreme market movements, the ranking of the models in Table 2.3 for 1985-2000 may be sensitive to the number of significant stock market downturns in this period. The results for the 1990-2000 period provide a check on this as there are only two days with $|R_t| > 5.0$ (R_t =-7.4, 27/10/97 and R_t =-6.5, 31/8/98) in this shorter period, while there are 6 days in 1985-90. The results in Table 2.3 reverse our ranking of the ARJI-R^2_{t-1} and ARJI-h_t models compared to the in-sample log-likelihood values.

Estimation of our dynamic jump models remains tractable by assuming the updating scheme for the conditional jump intensity can be projected onto past observables. As a result, *ex post* inference regarding the jump distribution is directly available from estimation. Figure 2.7 displays the *ex ante* and *ex post* probabilities of at least one jump occurring, calculated using Equations 2.2 and 2.6 for the ARJI-R^2_{t-1} model.

Figure 2.8 presents further evidence, in addition to Figure 2.4, supporting the hypothesis that the ARJI model may contain predictive information for stock market crashes. The graph in the top panel plots the market returns during the month of October 1987 and the negative 25% crash on the 19th. The predicted conditional average number of jumps rises from below .2 at the start of October 87 until it reaches 1.37 the day of the crash. Recall that this is an out-of-sample prediction, and λ_t is based on the previous day's information set. The value 1.37 represents over an eight fold increase in the conditional intensity compared to its unconditional value of 0.1547.

2.7.3 Comparison with Competing Models

Table 2.7 presents the results from different competing models. The simple GARCH(1,1) is also estimated and the results are reported in the first column. Note that there is strong presistence in the conditional variance as $\alpha + \beta = 0.9931$. The log likelihood value is relatively lower than those in jump models. This result confirms the importance of jumps in modeling this stock index. The Student-t distribution is reported in the second column which has essentially the same variance estimates. The degree of freedom parameter is 0.1623 which show that normal distribution is not an adequate

Table 2.4: Out-of-sample forecasting performance from different models, R^2

Sample period	GARCH	GARCH-t	G-Charlier	Exp. Power	Gen-t
1985-2000	0.269	0.266	0.272	0.271	0.267
1990-2000	0.334	0.335	0.335	0.336	0.336

This table shows the R^2 from a regression of the range statistic on a constant and the out-of-sample, one-period ahead forecast of the conditional standard deviation for the respective models.

representation of this particular dataset. A fat tail distribution like Student t can produce a much better likelihood value by capturing the excess kurtosis.

The results from the other three flexible distributions: Gram Charlier, Exponential Power, and Generalized t, are very similar in terms of the variance structure. The Gram Charlier distribution shows that the returns exhibit negative skewness and the Exponential Power distribution finds the dataset to be best represented by a distribution between the normal and the double exponential. The Generalized t distribution on the other hand finds the dataset does exhibit thick tails and the likelihood value is very close to the Student t distribution. In summary, the Student t distribution seems to be adequate in modelling the return series compared to the other three alternatives. The out-of-sample analysis is reported in Table 2.4. The GARCH model with Student t distribution again dominates all the alternatives. However, comparing these results to the one from the ARJI models has shown that none of these competing models can out perform the ARJI models regardless of whether or not the sample includes the 87 crashes.

Finally, Figure 2.9 shows conditional skewness and conditional kurtosis for the 1990s. Note that conditional skewness is usually negative, but can temporarily become positive after a sharp market drop. This implies that after a drop in the market, an improved daily return is the most likely occurrence.

2.8 Discussion

All models of conditional jump dynamics estimated on the DJIA returns show significant persistence in the conditional intensity. This indicates that the risk associated with jumps in stock market returns is systematic and should be important for derivative pricing. Furthermore, the jump dynamics do not reduce the GARCH effects in any of the models. The GARCH parameters across all models, including a no jump AR(2)-GARCH(1,1) model (not reported), are very similar. The ARJI family of models we investigate appear to be explaining dynamics not captured in the constant intensity specification.

Evidence of persistence in the jump intensity is also found in the analysis of Lin, Knight, and Satchell (1999) applied to intra-day equity returns. Lin, Knight, and Satchell propose a pure jump diffusion model that allows the intensity to depend upon the past conditional variance. Similar to our specification this model can capture time-dependence in the conditional jump intensity; however, the loglikelihood function does not have a closed form solution.

In the last section we found a robust result with respect to a switch in the jump direction after a stock market decrease. Both models with jump size dynamics have a significant $\theta_2 < 0$. Using the estimates from Table 2.6 for the ARJI-R_{t-1}^2 model, any decrease in the market of 2.5 percent or more, implies a positive conditional mean in next period's jump size distribution. Therefore, after a stock market crash, the likelihood of a jump next period does not necessarily decrease, but the likelihood of a negative jump decreases and the likelihood of a positive jump increases. This asymmetry in the jump direction is also found in the out-of-sample data period. For instance, there are 45 times in which returns decrease by 2.5% or more. Of the 45 times that $R_{t-1} < -2.5$, in 31 cases $R_t > 0$ in the next trading day. Conditional on $R_{t-1} < -2.5$, the samples averages of R_t, θ_t and λ_t (next period values) are 0.79, 0.31, and 0.90 respectively. This means that, on average, after a market downturn of -2.5% or more, a positive jump in returns is very likely to occur in the next trading day.

Previous research (Bates (2000), and Chernov, Gallant, Ghysels, and Tauchen (1999)) has investigated whether there is a relationship between the jump intensity and a stochastic volatility specification of volatility. In our model, the analogous relationship is between the conditional jump intensity parameterization and the GARCH specification. In general, we found mixed results. After permitting the variance of the jump size distribution to be a function of the GARCH variance (ARJI-h_t) we found no evidence (estimates not reported) that the GARCH process affects the conditional intensity specification for our data set. However, we did find that the GARCH variance was positive and significant in affecting λ_t in the ARJI-R_{t-1}^2 model.

This study documents evidence of significant conditional dynamics in the distribution governing the number of jumps and the jump size. An important topic from a risk management perspective is the prediction of extreme volatility. Our model captures this through jumps. In particular, we present evidence that before both the 1929 (in-sample) and 1987 (out-of-sample) stock market crashes the conditional expected number of jumps rises. This suggests that time-series data alone may contain predictive content that jump models such as those explored in this chapter could exploit in forecasting future market downturns. One way that jump predictability may work is that proportionately, jumps may become more important just before a crash. Thus, the GARCH or SV volatility component of the conditional variance may become less important in describing the total volatility while jumps become more important.

2.9 Conclusion

This chapter proposes an autoregressive conditional jump intensity (ARJI) model to capture jump dynamics in stock market returns. Our model extends the GARCH-constant jump intensity model with a time-varying jump intensity and jump size distribution. This chapter proposes a simple filter to provide inference regarding the number of jumps. The jump intensity is modeled as a parsimonious ARMA structure driven by an *ex post* measure of the jump probability.

We find significant time-variation in the conditional jump intensity and the jump size distribution in our application to daily stock market returns. Modeling jump dynamics in this chapter reveals several improvements compared to the basic GARCH-constant jump intensity model. First, the time-varying jump intensity provides good forecasts of stock market volatility. Second, incorporating the autoregressive jump intensity exploits additional structure ignored in a constant intensity model. Third, the ARJI model has significantly higher *ex ante* probabilities regarding jumps on the days of stock market downturns. Finally, we find evidence of an asymmetric jump direction around stock market crashes.

2.10 Appendix A

- Derivations of Equation (2.18)

$$\begin{aligned}
E[C_t^2|n_t = j, \Phi_{t-1}] &= \mathrm{Var}(C_t|n_t = j, \Phi_{t-1}) + E[C_t|n_t = j, \Phi_{t-1}]^2 \\
&= j\delta_t^2 + [j\theta_t]^2 \\
&= j\delta_t^2 + j^2\theta_t^2 \\
&= j\delta_t^2 + j\theta_t^2 + j^2\theta_t^2 - j\theta_t^2 \\
&= j(\delta_t^2 + \theta_t^2) + \theta_t^2(j^2 - j)
\end{aligned}$$

$$\begin{aligned}
E[C_t^2|\Phi_{t-1}] &= \sum_{j=0}^{\infty} E[C_t^2|n_t = j, \Phi_{t-1}]P(n_t = j|\Phi_{t-1}) \\
E[C_t^2|\Phi_{t-1}] &= \sum_{j=0}^{\infty}(j(\delta_t^2 + \theta_t^2) + \theta_t^2(j^2 - j))P(n_t = j|\Phi_{t-1}) \\
&= (\delta_t^2 + \theta_t^2)\sum_{j=0}^{\infty} jP(n_t = j|\Phi_{t-1}) + \theta_t^2\sum_{j=0}^{\infty}(j^2 - j)P(n_t = j|\Phi_{t-1}) \\
&= (\delta_t^2 + \theta_t^2)\sum_{j=0}^{\infty} jP(n_t = j|\Phi_{t-1}) + \theta_t^2\sum_{j=0}^{\infty} j^2P(n_t = j|\Phi_{t-1}) \\
&\quad -\theta_t^2\sum_{j=0}^{\infty} jP(n_t = j|\Phi_{t-1}) \\
&= (\delta_t^2 + \theta_t^2)\sum_{j=0}^{\infty} jP(n_t = j|\Phi_{t-1}) + \theta_t^2(\lambda_t(1 + \lambda_t)) - \theta_t^2\lambda_t \\
&= (\delta_t^2 + \theta_t^2)\lambda_t + \theta_t^2\lambda_t^2
\end{aligned}$$

- Derivations of Equation(2.21)

$$\text{Skew}(R_t|\Phi_{t-1}) = \frac{\mu_R^3}{\sigma_R^3}$$

where μ_R^3 is the 3rd moment around the mean for R_t and σ_R is the standard deviation of the returns. Assuming $\mu = \phi_1 = \phi_2 = 0$ and using the fact that skewness of normal error terms is zero, the third moment around the mean for R_t is defined by

$$\begin{aligned}
\mu_R^3 &= E[C - E[C|\Phi_{t-1}]|\Phi_{t-1}]^3 \\
&= E[C^3|\Phi_{t-1}] - 3E[C|\Phi_{t-1}]\text{Var}(C|\Phi_{t-1}) - [E[C|\Phi_{t-1}]]^3
\end{aligned}$$

The only unknown in the above equation is $E[C^3|\Phi_{t-1}]$ which can be derived using equation 2.16

$$E[C_t^3|\Phi_{t-1}] = \sum_{j=0}^{\infty} E[C_t^3|n_t = j, \Phi_{t-1}]P(n_t = j|\Phi_{t-1}), \quad i > 0.$$

and

$$E[C_t^3|n_t = j, \Phi_{t-1}] = E[C|n_t = j, \Phi_{t-1}]((E[C|n_t = j, \Phi_{t-1}])^2$$

$$
\begin{aligned}
& +3\mathrm{Var}(C|n_t = j, \Phi_{t-1})) \\
&= j\theta_t(j^2\theta_t^2 + 3j\delta_t^2) \\
&= j^3\theta_t^3 + 3j^2\theta_t\delta_t^2
\end{aligned}
$$

$$
\begin{aligned}
E[C_t^3|\Phi_{t-1}] &= \sum_{j=0}^{\infty}(j^3\theta_t^3 + 3j^2\theta_t\delta_t^2)P(n_t = j|\Phi_{t-1}) \\
&= \sum_{j=0}^{\infty} j^3\theta_t^3 P(n_t = j|\Phi_{t-1}) + \sum_{j=0}^{\infty} 3j^2\theta_t\delta_t^2 P(n_t = j|\Phi_{t-1}) \\
&= \theta_t^3 P\sum_{j=0}^{\infty} j^3 P(n_t = j|\Phi_{t-1}) + 3\theta_t\delta_t^2\sum_{j=0}^{\infty} j^2 P(n_t = j|\Phi_{t-1}) \\
&= \theta_t^3(\lambda_t(1 + 3\lambda_t + \lambda_t^2)) + 3\theta_t\delta_t^2\lambda_t(1+\lambda_t) \\
&= \theta_t^3\lambda_t + 3\theta_t^3\lambda_t^2 + \theta_t^3\lambda_t^3 + 3\theta_t\delta_t^2\lambda_t + 3\theta_t\delta_t^2\lambda_t^2
\end{aligned}
$$

Substituting this into μ_R^3 gives

$$
\begin{aligned}
\mu_R^3 &= E[C^3|\Phi_{t-1}] - 3E[C|\Phi_{t-1}]\mathrm{Var}(C|\Phi_{t-1}) - [E[C|\Phi_{t-1}]]^3 \\
&= \theta_t^3\lambda_t + 3\theta_t^3\lambda_t^2 + \theta_t^3\lambda_t^3 + 3\theta_t\delta_t^2\lambda_t + 3\theta_t\delta_t^2\lambda_t^2 - 3\theta_t\lambda_t(\delta_t^2\lambda_t + \theta_t^2\lambda_t) - \theta_t^3\lambda_t^3 \\
&= \theta_t^3\lambda_t + 3\theta_t\delta_t^2\lambda_t \\
&= \lambda_t(\theta_t^3 + 3\theta_t\delta_t^2)
\end{aligned}
$$

Finally, the conditional skewness can be derived as

$$
\begin{aligned}
\mathrm{Skew}(R_t|\Phi_{t-1}) &= \frac{\mu_R^3}{\sigma_R^3} \\
&= \frac{\lambda_t(\theta_t^3 + 3\theta_t\delta_t^2)}{(h_t + \lambda_t\delta_t^2 + \lambda_t\theta_t^2)^{3/2}}
\end{aligned}
$$

- Derivations of Equation(2.22)

  We can derive the conditional kurtosis in a similar fashion.

$$
\mathrm{Kurt}(R_t|\Phi_{t-1}) = \frac{\mu_R^4}{\sigma_R^4}
$$

where μ_R^4 is the 4th moment around the mean for R_t and σ_R is the standard deviation of the returns. Therefore the fourth moment around the mean for R_t is

$$\begin{aligned}
\mu_R^4 &= E[\sqrt{h_t}z_t + C - E[C|\Phi_{t-1}]|\Phi_{t-1}]^4 \\
&= E[(\sqrt{h_t}z_t + C)^4|\Phi_{t-1}] - 4E[C|\Phi_{t-1}]E[(\sqrt{h_t}z_t + C)^3|\Phi_{t-1}] \\
&\quad +6(E[C|\Phi_{t-1}])^2E[(\sqrt{h_t}z_t + C)^2|\Phi_{t-1}] - 3(E[C|\Phi_{t-1}])^4
\end{aligned}$$

The only unknown in the above equation is $E[(\sqrt{h_t}z_t + C)^4|\Phi_{t-1}]$ which can be derived using equation 2.16

$$E[C_t^4|\Phi_{t-1}] = \sum_{j=0}^{\infty} E[C_t^4|n_t = j, \Phi_{t-1}]P(n_t = j|\Phi_{t-1}), \quad i > 0.$$

and

$$\begin{aligned}
E[C_t^4|n_t = j, \Phi_{t-1}] &= E[C|n_t = j, \Phi_{t-1}]^4 \\
&\quad +6(E[C|n_t = j, \Phi_{t-1}])^2\mathrm{Var}(C|n_t = j, \Phi_{t-1}) \\
&\quad +3(\mathrm{Var}(C|n_t = j, \Phi_{t-1}))^2 \\
&= j^4\theta_t^4 + 6j^2\theta_t^2(j\delta_t^2) + 3j^2\delta_t^4 \\
&= j^4\theta_t^4 + 6j^3\theta_t^2\delta_t^2 + 3j^2\delta_t^4 \\
E[C_t^4|\Phi_{t-1}] &= \sum_{j=0}^{\infty}(j^4\theta_t^4 + 6j^3\theta_t^2\delta_t^2 + 3j^2\delta_t^4)P(n_t = j|\Phi_{t-1}) \\
&= \sum_{j=0}^{\infty} j^4\theta_t^4 P(n_t = j|\Phi_{t-1}) + \sum_{j=0}^{\infty} 6j^3\theta_t^2\delta_t^2 P(n_t = j|\Phi_{t-1}) \\
&\quad +\sum_{j=0}^{\infty} 3j^2\delta_t^4 P(n_t = j|\Phi_{t-1}) \\
&= \theta_t^4 \sum_{j=0}^{\infty} j^4 P(n_t = j|\Phi_{t-1}) + 6\theta_t^2\delta_t^2 \sum_{j=0}^{\infty} j^3 P(n_t = j|\Phi_{t-1}) \\
&\quad +3\delta_t^4 \sum_{j=0}^{\infty} j^2 P(n_t = j|\Phi_{t-1}) \\
&= \theta_t^4\lambda_t(1 + 7\lambda_t + 6\lambda_t^2 + \lambda_t^3) + 6\theta_t^2\delta_t^2(\lambda_t(1 + 3\lambda_t + \lambda_t^2))
\end{aligned}$$

$$
\begin{aligned}
& +3\delta_t^4\lambda_t(1+\lambda_t) \\
= \; & \lambda_t(\theta_t^4 + 6\theta_t^2\delta_t^2 + 3\delta_t^4) + \lambda_t^2(7\theta_t^4 + 18\theta_t^2\delta_t^2 + 3\delta_t^4) \\
& +\lambda_t^3(6\theta_t^4 + 6\theta_t^2\delta_t^2) + \lambda_t^4\theta_t^4
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
E(\sqrt{h_t}z_t + C)^4 &= E((\sqrt{h_t}z_t)^4) + 4E((\sqrt{h_t}z_t)^3)E(C) \\
&\quad +6E((\sqrt{h_t}z_t)^2)E(C^2) + 4E(\sqrt{h_t}z_t)E(C^3) + E(C^4) \\
&= 3h_t^2 + 6h_t((\delta_t^2 + \theta_t^2)\lambda_t + \theta_t^2\lambda_t^2) \\
&\quad +\lambda_t(\theta_t^4 + 6\theta_t^2\delta_t^2 + 3\delta_t^4) + \lambda_t^2(7\theta_t^4 + 18\theta_t^2\delta_t^2 + 3\delta_t^4) \\
&\quad +\lambda_t^3(6\theta_t^4 + 6\theta_t^2\delta_t^2) + \lambda_t^4\theta_t^4
\end{aligned}
$$

Substituting this into μ_R^4 gives

$$
\begin{aligned}
\mu_R^4 &= E[\sqrt{h_t}z_t + C - E[C|\Phi_{t-1}]|\Phi_{t-1}]^4 \\
&= E[(\sqrt{h_t}z_t + C)^4|\Phi_{t-1}] - 4E[C|\Phi_{t-1}]E[(\sqrt{h_t}z_t + C)^3|\Phi_{t-1}] \\
&\quad +6(E[C|\Phi_{t-1}])^2E[(\sqrt{h_t}z_t + C)^2|\Phi_{t-1}] - 3(E[C|\Phi_{t-1}])^4 \\
&= 3h_t^2 + 6h_t((\delta_t^2 + \theta_t^2)\lambda_t + \theta_t^2\lambda_t^2) \\
&\quad +\lambda_t(\theta_t^4 + 6\theta_t^2\delta_t^2 + 3\delta_t^4) + \lambda_t^2(7\theta_t^4 + 18\theta_t^2\delta_t^2 + 3\delta_t^4) \\
&\quad +\lambda_t^3(6\theta_t^4 + 6\theta_t^2\delta_t^2) + \lambda_t^4\theta_t^4 \\
&\quad -4\theta_t\lambda_t(3h_t\lambda_t\theta_t + \theta_t^3\lambda_t + 3\theta_t^3\lambda_t^2 + \theta_t^3\lambda_t^3 + 3\theta_t\delta_t^2\lambda_t) \\
&\quad +6(\lambda_t^2\theta_t^2(h_t + (\delta_t^2 + \theta_t^2)\lambda_t + \theta_t^2\lambda_t^2) - 3\lambda_t^4\theta_t^4 \\
&= 3h_t^2 + 6h_t((\delta_t^2 + \theta_t^2)\lambda_t) + 3((\delta_t^2 + \theta_t^2)\lambda_t)^2 + \lambda_t(\theta_t^4 + 6\theta_t^2\delta_t^2 + 3\delta_t^4) \\
&= 3(h_t + \lambda_t(\delta_t^2 + \theta_t^2))^2 + \lambda_t(\theta_t^4 + 6\theta_t^2\delta_t^2 + 3\delta_t^4)
\end{aligned}
$$

Finally, the conditional kurtosis can be derived as

$$
\mathrm{Kurt}(R_t|\Phi_{t-1}) = \frac{\mu_R^4}{\sigma_R^4}
$$

$$= \quad 3 + \frac{\lambda_t(\theta_t^4 + 6\theta_t^2\delta_t^2 + 3\delta_t^4)}{(h_t + \lambda_t\delta_t^2 + \lambda_t\theta_t^2)^2}$$

- Properties of Normal Distribution

The normal probability density function with mean μ and variance σ^2 is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp^{-(x-\mu)^2/2\sigma^2}.$$

The Characteristic Function is

$$\Phi(t) = \exp^{imt-\sigma^2t^2/2}$$

and the Moment-Generating function is

$$M(t) = \exp^{\mu t+\sigma^2t^2/2}.$$

The Cumulant Generating Function is

$$K(h) = ln(\exp^{\lambda h}\exp^{\sigma^2h^2/2}).$$

The first four moments around zero are

$$\begin{aligned} \mu_1' &= \mu \\ \mu_2' &= \mu^2 + \sigma^2 \\ \mu_3' &= \mu(\mu^2 + 3\sigma^2) \\ \mu_4' &= \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4 \end{aligned}$$

The first four moments around the mean are

$$\begin{aligned} \mu_1 &= \mu \\ \mu_2 &= \sigma^2 \\ \mu_3 &= 0 \\ \mu_4 &= 3\sigma^4 \end{aligned}$$

- Properties of Poisson Distribution

The Poisson density function with intensity parameter λ is defined by

$$P(X=j)=\frac{\exp(-\lambda)\lambda^j}{j!}.$$

The Characteristic function is

$$\Phi(t)=\exp^{m(\exp^{it}-1)}$$

and the Moment Generating function is

$$M(t)=\exp^{\lambda(\exp^{t}-1)}.$$

The Cumulant Generating function is

$$K(h)=\lambda(\exp^h-1).$$

The first four moments around zero are

$$\begin{aligned}
\mu_1' &= \lambda \\
\mu_2' &= \lambda(1+\lambda) \\
\mu_3' &= \lambda(1+3\lambda+\lambda^2) \\
\mu_4' &= \lambda(1+7\lambda+6\lambda^2+\lambda^3)
\end{aligned}$$

The first four moments around the mean are

$$\begin{aligned}
\mu_1 &= \lambda \\
\mu_2 &= \lambda \\
\mu_3 &= \lambda \\
\mu_4 &= \lambda(1+3\lambda)
\end{aligned}$$

Table 2.5: Estimates of the Constant Intensity Jump Model for different sample periods

$$R_t = \mu + \sum_{i=1}^{2} \phi_i R_{t-i} + \sqrt{h_t} z_t + \sum_{k=1}^{n_t} Y_{t,k}$$

$$Y_{t,k} \sim NID(\theta, \delta^2), \quad h_t = \omega + \alpha \epsilon_{t-1}^2 + \beta h_{t-1}$$

$$P(n_t = j) = \frac{\exp^{-\lambda} \lambda^j}{j!}, \quad z_t \sim NID(0,1)$$

Parameter	1928-1950	1951-1969	1970-1984
μ	0.0753	0.2138	-0.0025
	(0.0097)	(0.0331)	(0.0142)
ϕ_1	0.0940	0.1848	0.1623
	(0.0125)	(0.0154)	(0.0166)
ϕ_2	-0.0627	-0.0818	-0.0296
	(0.0123)	(0.0151)	(0.0166)
ω	0.0009	2.07E-13	0.0072
	(0.0009)	(0.0005)	(0.0024)
α	0.0596	0.1118	0.0402
	(0.0058)	(0.0139)	(0.0056)
β	0.9213	0.8008	0.9472
	(0.0069)	(0.0264)	(0.0074)
δ	1.4624	0.2711	1.430
	(0.1345)	(0.0346)	(0.2789)
θ	-0.6447	-0.1129	0.9882
	(0.1147)	(0.0184)	(0.7942)
λ	0.1116	1.6742	0.0196
	(0.0186)	(0.4043)	(0.0166)
Q^2	52.72	27.25	19.01
	[0.00]	[0.03]	[0.21]
Q_{ξ_t}	23.74	23.66	30.14
	[0.07]	[0.07]	[0.01]
log likelihood	-9149.74	-4213.27	-4814.20

Standard errors are in parentheses. p-values are in square brackets. Q^2 is the modified Ljung-Box Portmanteau test, robust to heteroskedasticity, for serial correlation in the squared standardized residuals with 15 lags for the respective models. Q_{ξ_t} is the same test for serial correlation in the jump intensity residuals.

Table 2.6: Estimates of the ARJI Models, 1928-84

$$R_t = \mu + \sum_{i=1}^{2} \phi_i R_{t-i} + \sqrt{h_t} z_t + \sum_{k=1}^{n_t} Y_{t,k}$$
$$Y_{t,k} \sim N(\theta_t, \delta_t^2), \quad h_t = \omega + \alpha\epsilon_{t-1}^2 + \beta h_{t-1}$$
$$\lambda_t = \lambda_0 + \rho\lambda_{t-1} + \gamma\xi_{t-1}$$
$$\eta_t = \eta_0 + \eta_1 R_{t-1} D(R_{t-1}) + \eta_2 R_{t-1}(1 - D(R_{t-1}))$$
$$\text{ARJI} - h_t : \delta_t^2 = \zeta_0^2 + \zeta_1 h_t, \quad \text{ARJI} - R_{t-1}^2 : \delta_t^2 = \zeta_0^2 + \zeta_1 R_{t-1}^2$$

Parameter	Constant	ARJI	ARJI-R_{t-1}^2	ARJI-h_t
μ	0.0692	0.0570	0.0713	0.1091
	(0.0083)	(0.0070)	(0.0079)	(0.0100)
ϕ_1	0.1389	0.1331	0.1486	0.1209
	(0.0084)	(0.0085)	(0.0099)	(0.0122)
ϕ_2	-0.0561	-0.0581	-0.0638	-0.0674
	(0.0083)	(0.0085)	(0.0085)	(0.0084)
ω	0.0005	0.0030	0.0020	0.0015
	(0.0008)	(0.0006)	(0.0005)	(0.0003)
α	0.0665	0.0346	0.0294	0.0201
	(0.0041)	(0.0041)	(0.0035)	(0.0029)
β	0.9192	0.9494	0.9564	0.9599
	(0.0047)	(0.0047)	(0.0041)	(0.0045)
ζ_0	0.8744	1.1779	0.8728	0.0000
	(0.0927)	(0.1075)	(0.0740)	(0.1445)
ζ_1			0.1337	1.6041
			(0.0440)	(0.1329)
η_0	-0.3222	-0.3962	-0.5218	-0.2955
	(0.0539)	(0.0626)	(0.0760)	(0.0428)
η_1			0.0322	0.1170
			(0.0503)	(0.0348)
η_2			-0.2069	-0.0820
			(0.0422)	(0.0292)
λ_0	0.1512	0.0131	0.0189	0.0695
	(0.0374)	(0.0032)	(0.0043)	(0.0157)
ρ		0.9153	0.9098	0.8412
		(0.0204)	(0.0196)	(0.0265)
γ		0.4919	0.5230	0.4461
		(0.0773)	(0.0801)	(0.0640)
Q^2	18.27	7.54	14.70	7.93
	[0.25]	[0.94]	[0.47]	[0.93]
Q_{ξ_t}	30.96	9.84	15.60	14.75
	[0.01]	[0.83]	[0.41]	[0.47]
log likelihood	-18315.61	-18276.10	-18232.45	-18143.40

Standard errors are in parentheses. p-values are in square brackets. Q^2 is the modified Ljung-Box Portmanteau test, robust to heteroskedasticity, for serial correlation in the squared standardized residuals with 15 lags for the respective models. Q_{ξ_t} is the same test for serial correlation in the jump intensity residuals. Constant is the constant jump intensity, constant jump size model.

Table 2.7: Estimates of Competing Models, 1928-84

Parameter	GARCH	GARCH-t	Gram-Charlier	Exp. Power	Generalized-t
μ	0.0304	0.0405	0.0244	0.0400	0.0405
	(0.0056)	(0.0053)	(0.0057)	(0.0053)	(0.0052)
ϕ_1	0.1445	0.1391	0.1348	0.1354	0.1385
	(0.0087)	(0.0082)	(0.0083)	(0.0080)	(0.0082)
ϕ_2	-0.0399	-0.0541	-0.0556	-0.0573	-0.0546
	(0.0087)	(0.0081)	(0.0083)	(0.0078)	(0.0081)
ω	0.0080	0.0062	0.0081	0.0069	0.0082
	(0.0009)	(0.0009)	(0.0010)	(0.0010)	(0.0012)
α	0.0784	0.0744	0.0807	0.0772	0.0993
	(0.0043)	(0.0052)	(0.0050)	(0.0052)	(0.0071)
β	0.9147	0.9211	0.9143	0.9168	0.9206
	(0.0045)	(0.0052)	(0.0052)	(0.0054)	(0.0052)
$1/v$		0.1623			
		(0.0078)			
λ_3			-0.16720		
			(0.0279)		
λ_4			1.0633		
			(0.0558)		
$1/\gamma$				1.3379	
				(0.0199)	
p					1.9075
					(0.0747)
q					3.5584
					(0.4818)
log likelihood	-18671.51	-18214.50	-18357.36	-18271.67	-18212.96

Standard errors are in parentheses. p-values are in square brackets. The Gram Charlier (Lee and Tse 1991) density is defined as

$$GC(u) = \phi(u)(1 + \lambda_3 H_3(u)/6 + \lambda_4 H_4(u)/24)$$
$$H_3(u) = u^3 - 3u \quad ; \quad H_4(u) = u^4 - 6u^3 + 3.$$

The Exponential Power (Box and Tiao 1973) density is defined as

$$EP(u) = w(u)exp[-c(\gamma)|u|^{1/\gamma}]$$
$$w(\gamma) = \frac{1}{2\gamma}\left[\frac{\Gamma(3\gamma)}{\Gamma^3(\gamma)}\right]^{1/2} \quad ; \quad c(\gamma) = \left[\frac{\Gamma(3\gamma)}{\Gamma(\gamma)}\right]^{\frac{1}{2\gamma}}.$$

Note that EP = Double Exponential when $1/\gamma = 1$; EP = Normal when $1/\gamma = 2.0$; EP = Uniform when $1/\gamma = \infty$. The Generalized t distribution (McDonald and Newey 1988) is defined by

$$GT(u_t) = \frac{p}{2h_t q^{1/p} B(1/p, q)\left(1 + \frac{|u_t|^p}{q h_t^p}\right)^{q+\frac{1}{p}}}$$

where B is the Beta function.

Figure 2.3: ARJI, In-Sample R_t, λ_t, and $\sqrt{\mathrm{Var}_{t-1}(R_t)}$

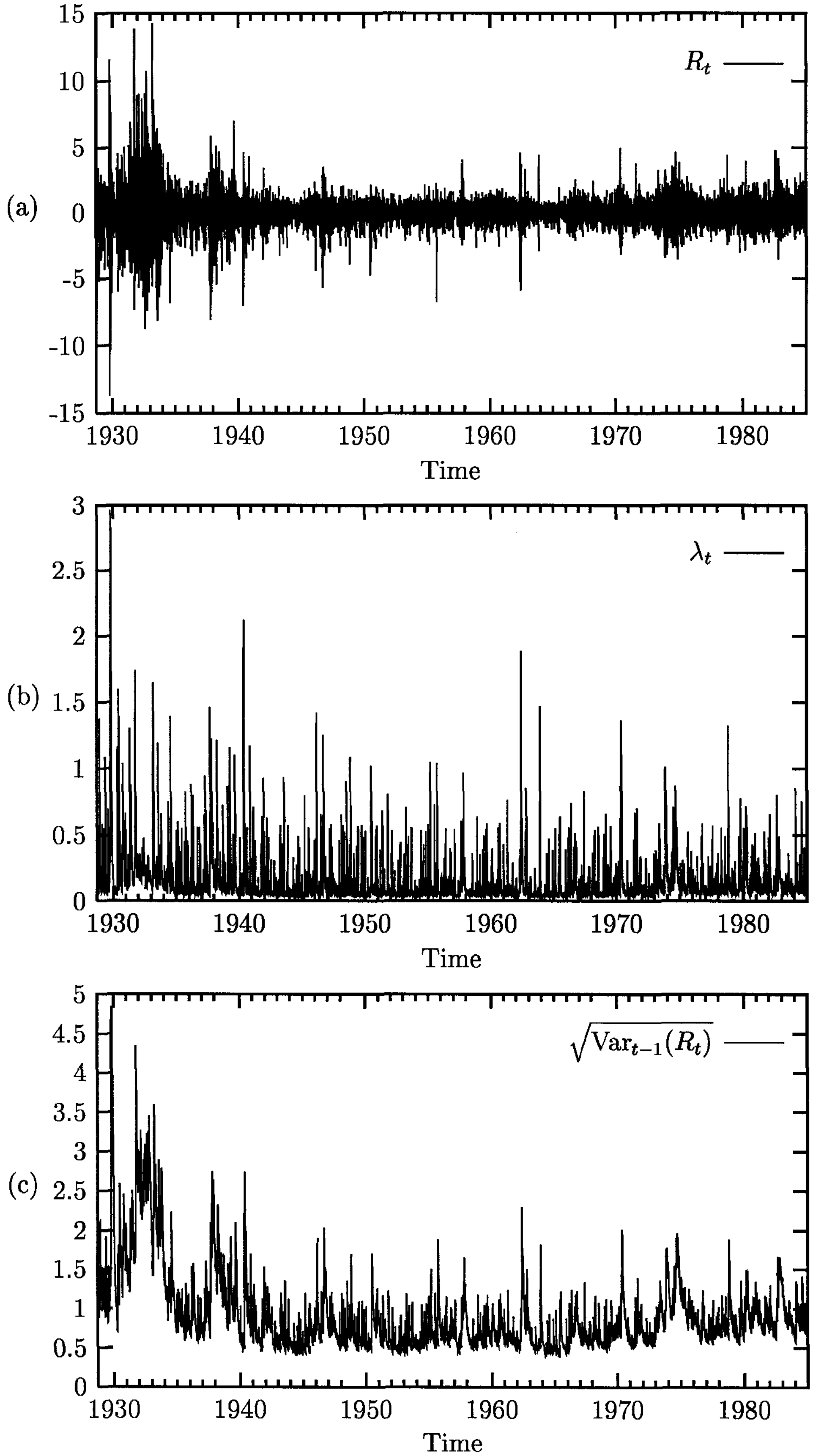


Figure 2.4: ARJI, In-sample R_t and λ_t, (Oct 15 - Oct 28, 1929)

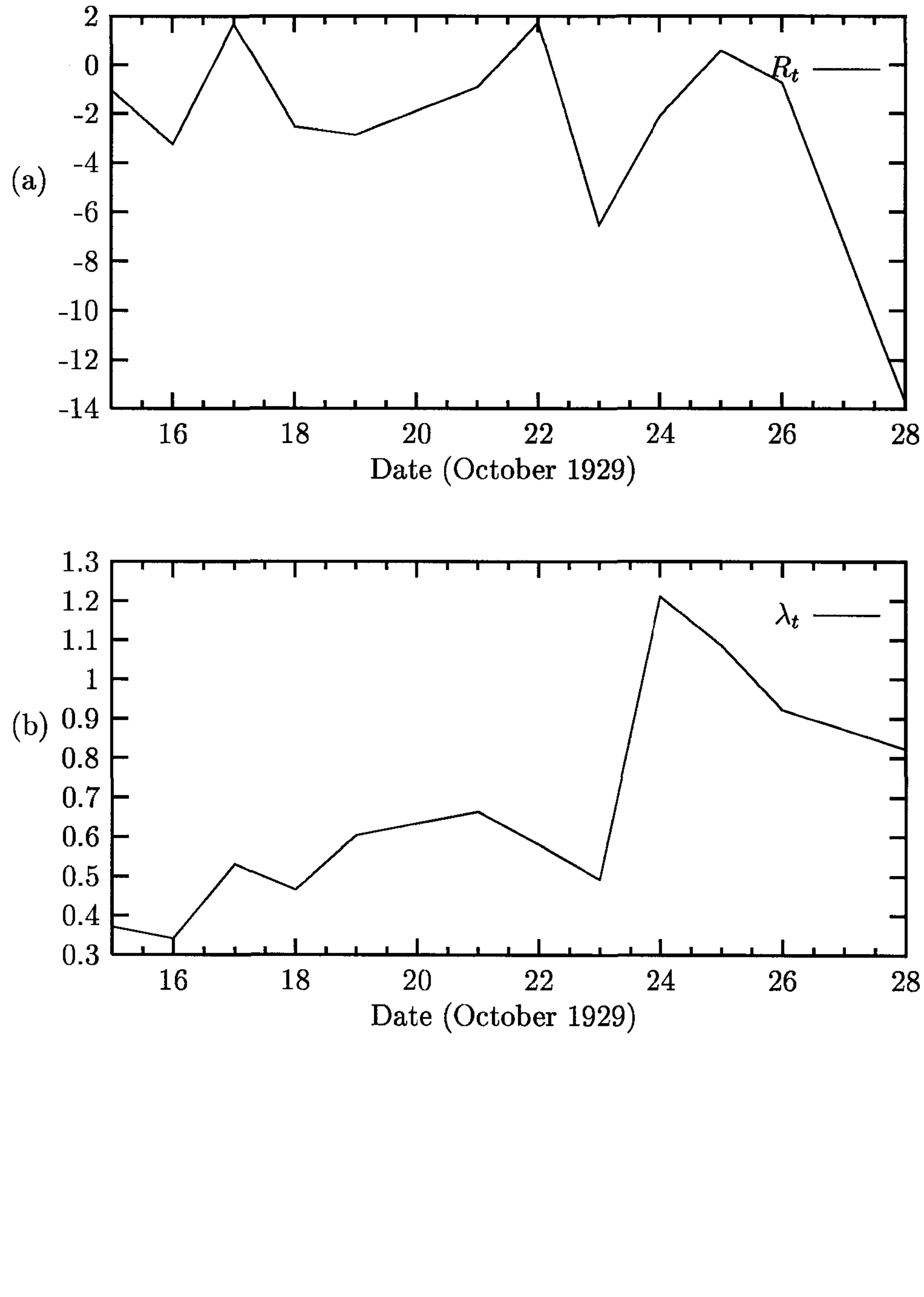


Figure 2.5: ARJI-h_t, In-sample conditional mean and variance from jumps

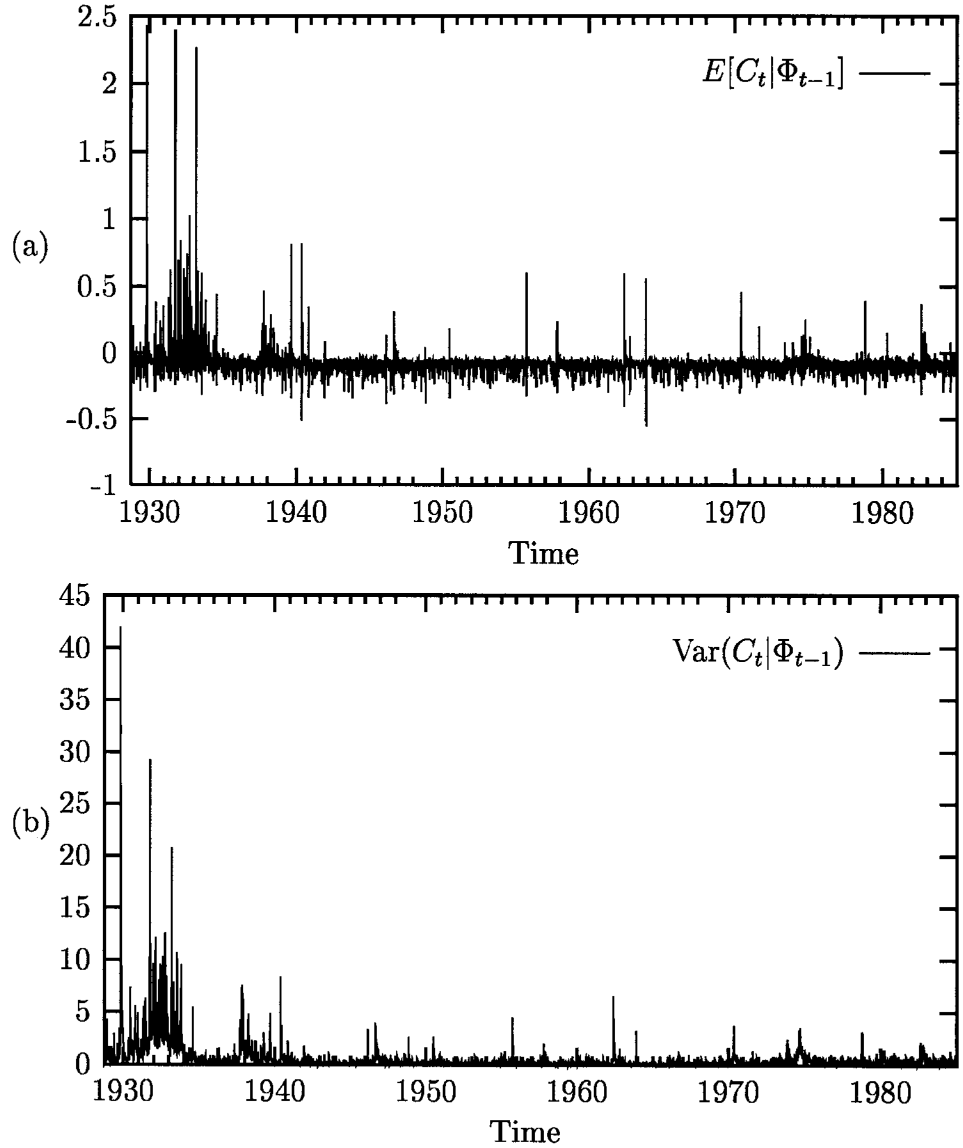


Figure 2.6: ARJI-h_t, Out-of-Sample R_t and λ_t

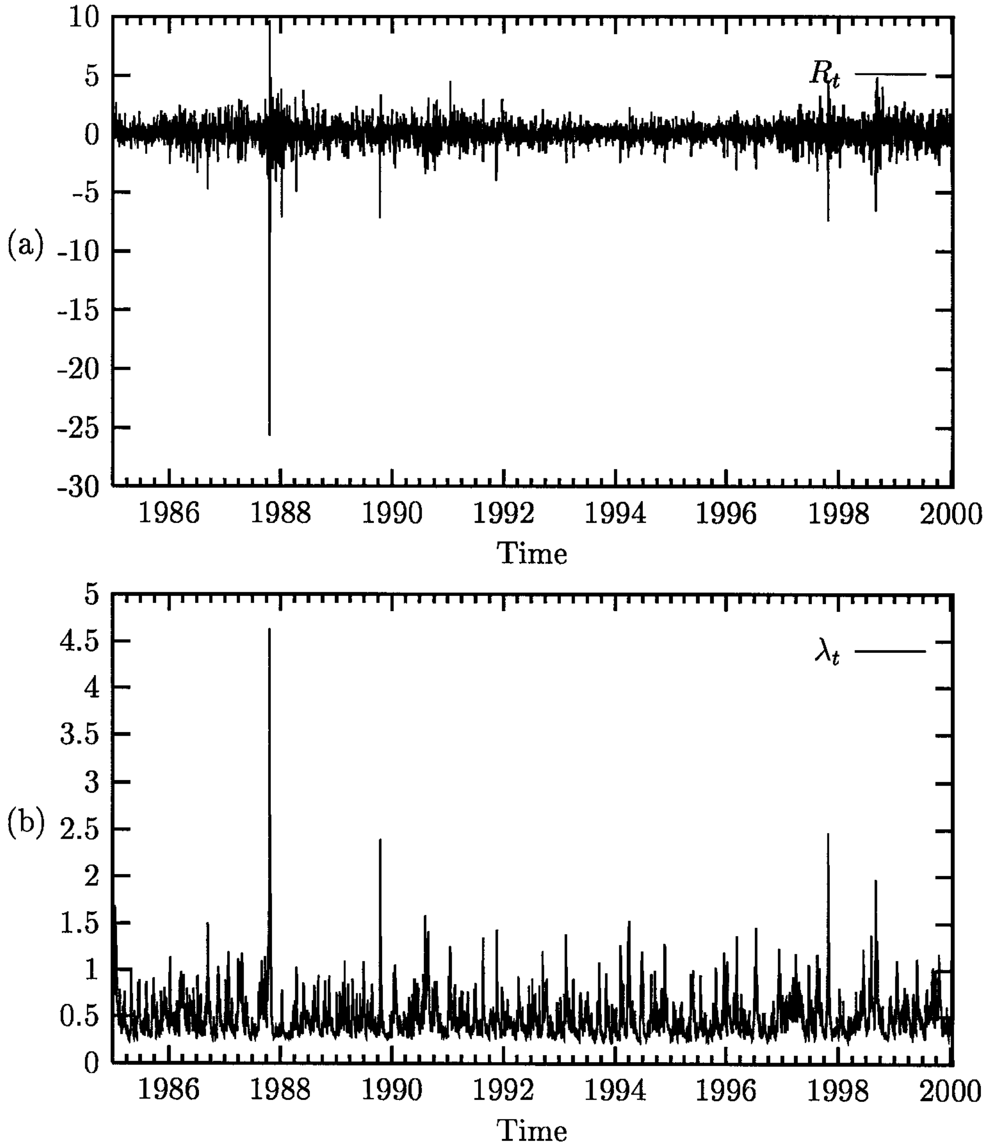


Figure 2.7: ARJI-R_{t-1}^2, Probability of at least one jump P(n_t ≥1)

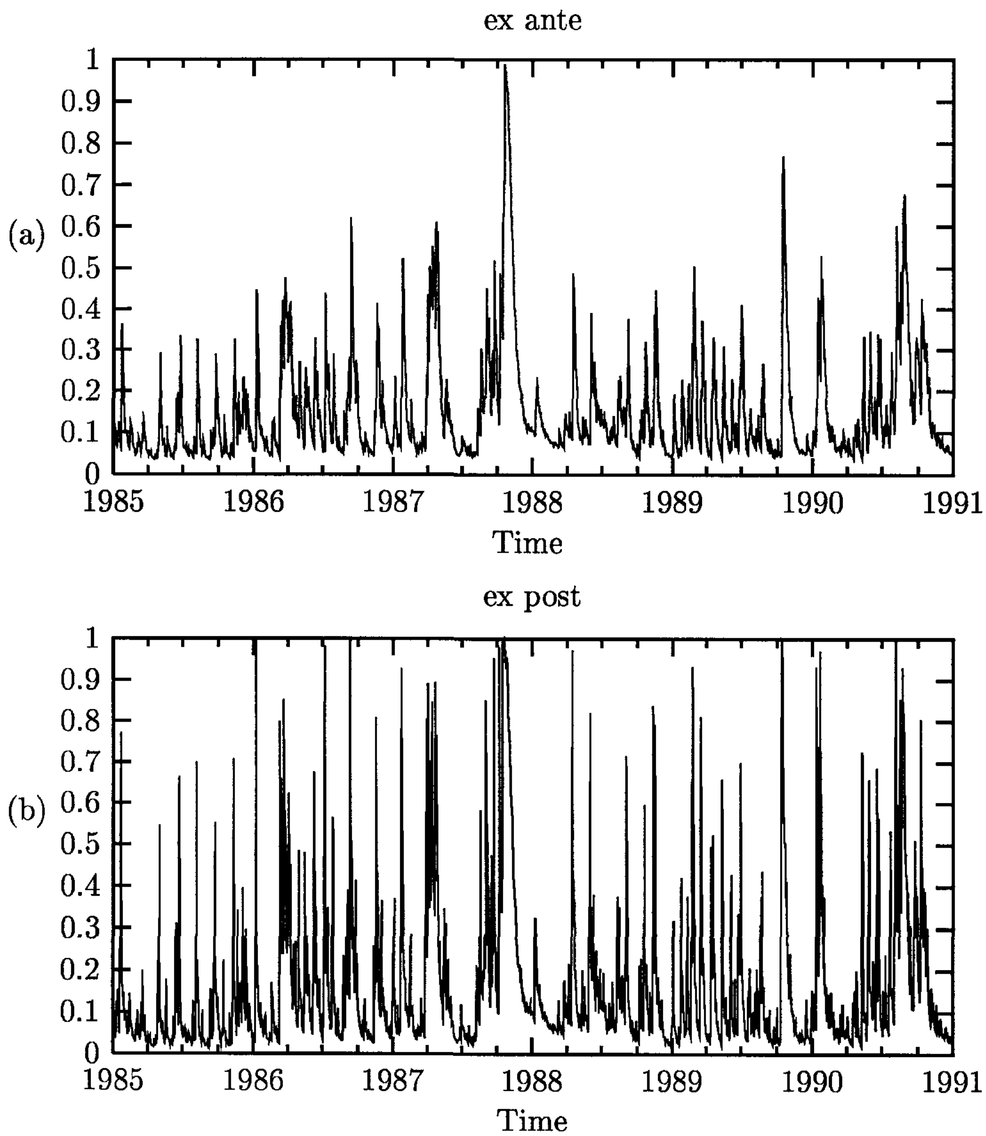


Figure 2.8: ARJI, Out-of-sample R_t and λ_t, (Oct 1 - Oct 19, 1987)

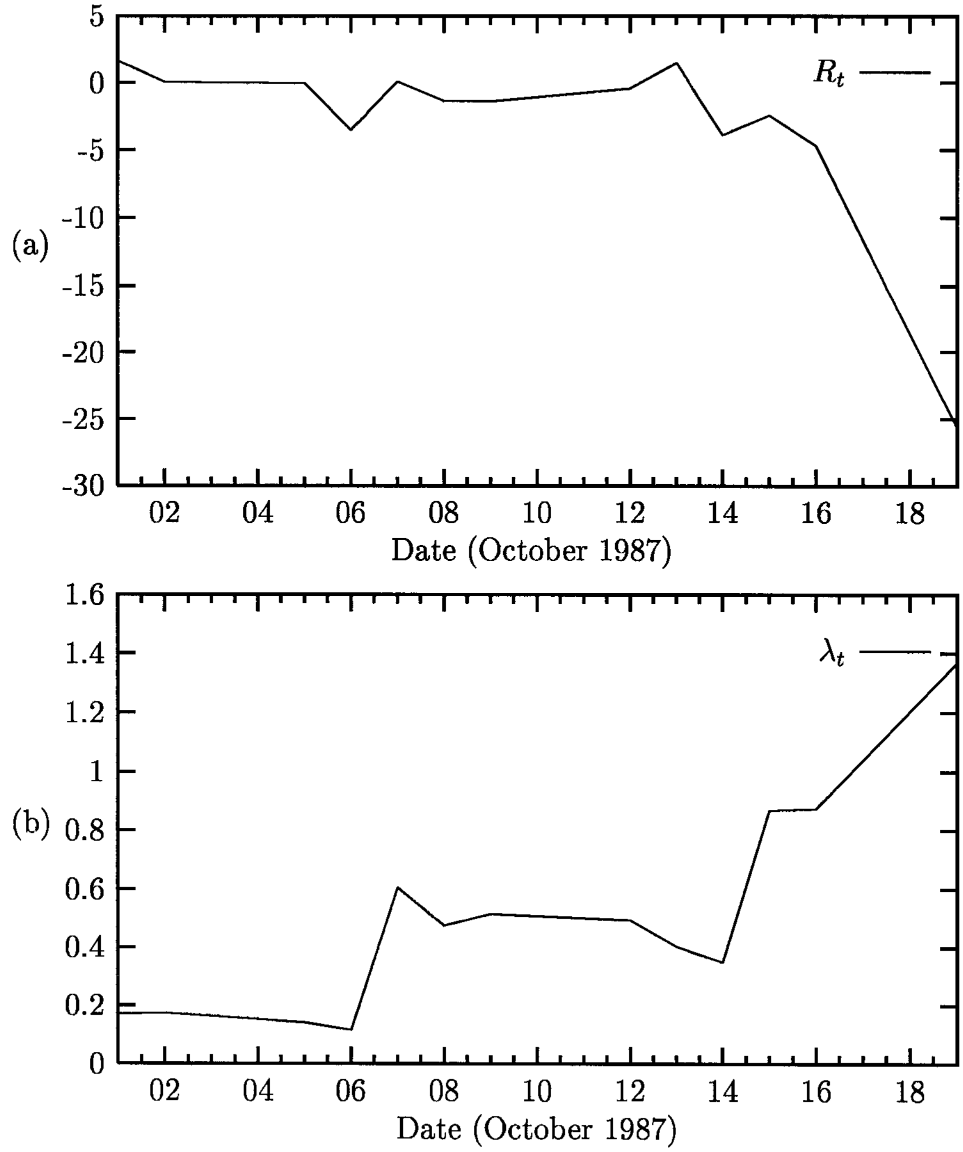


Figure 2.9: ARJI-h_t, Time-varying skewness and kurtosis

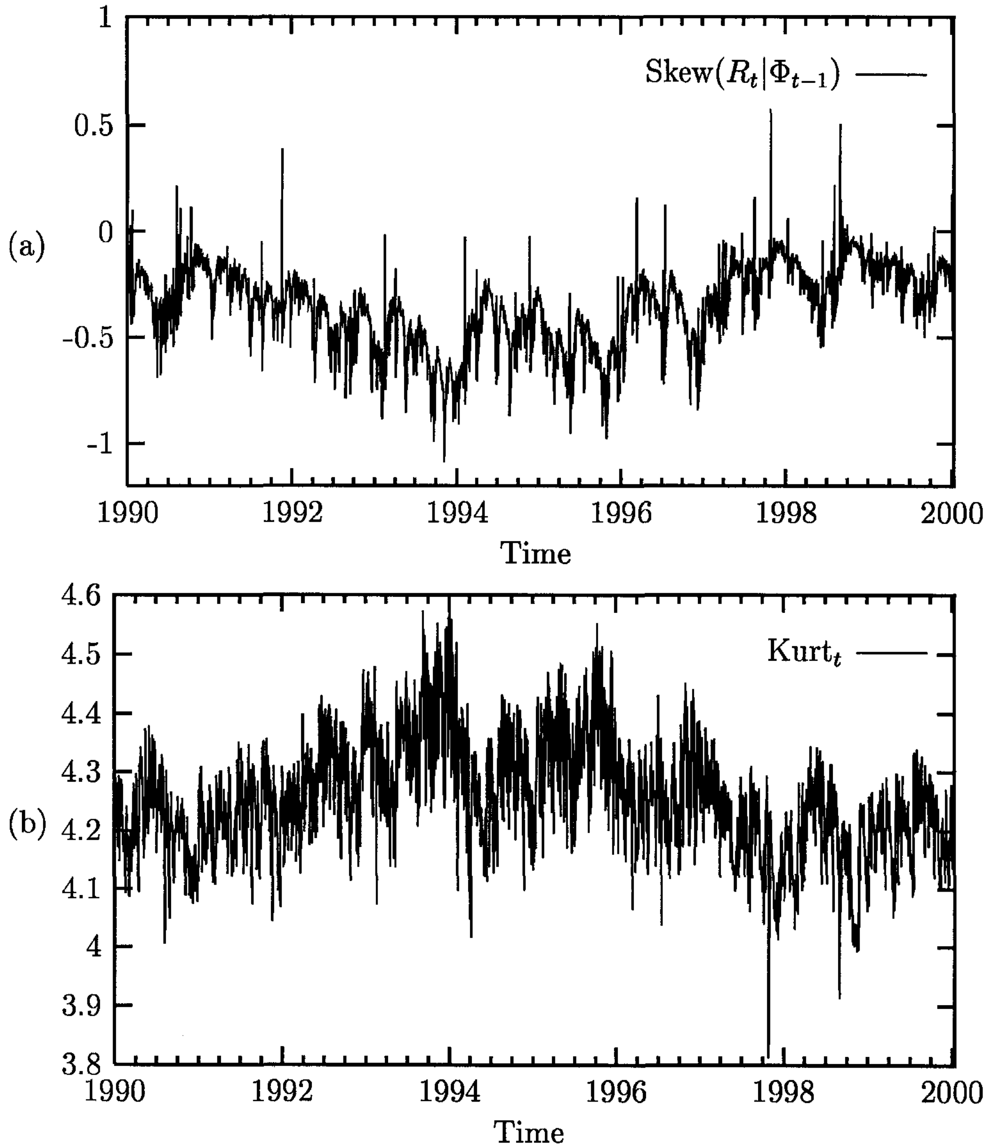
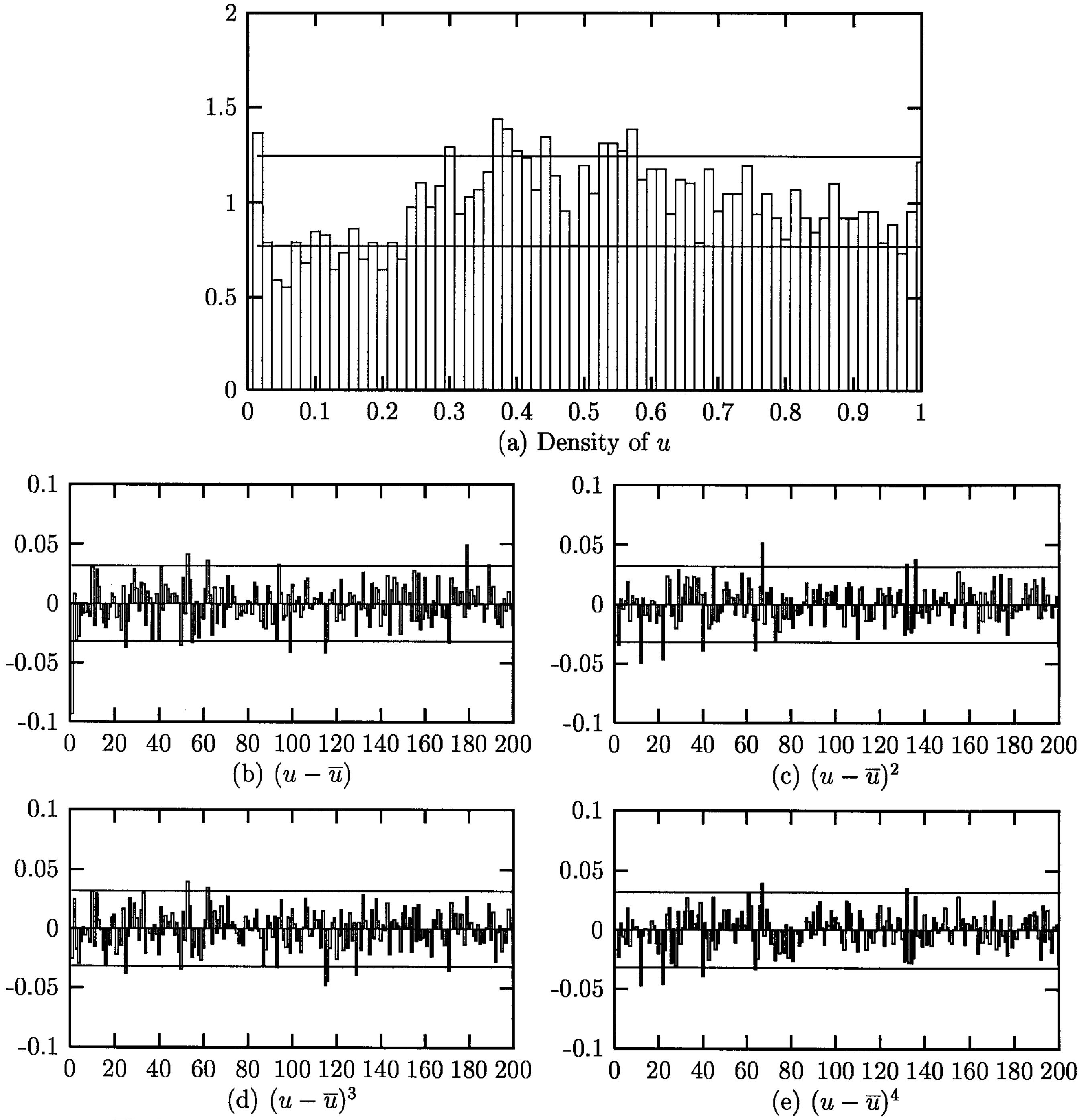
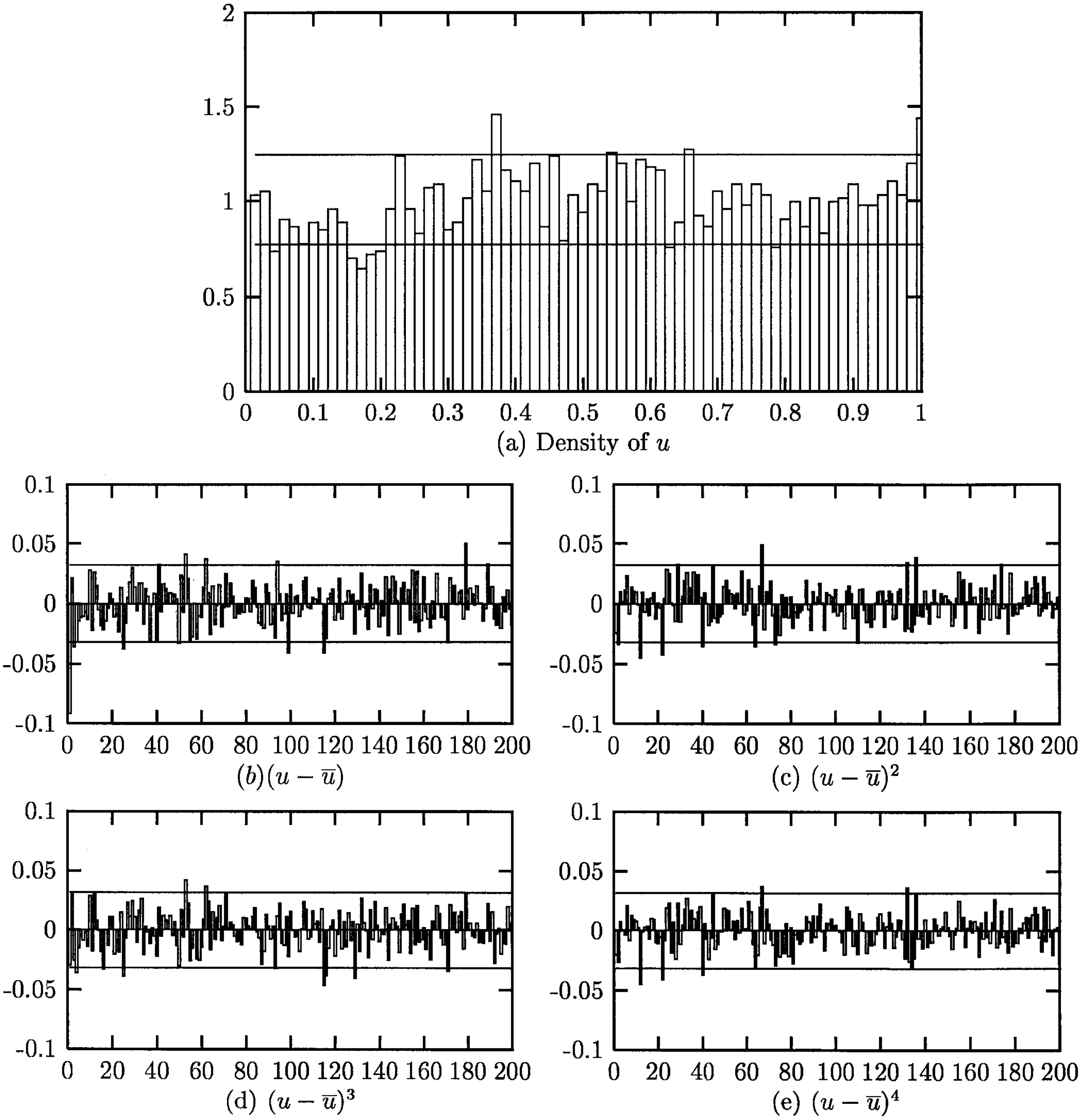


Figure 2.10: GARCH(1,1), Estimates of the density of u_t and autocorrelation functions of powers of u_t.



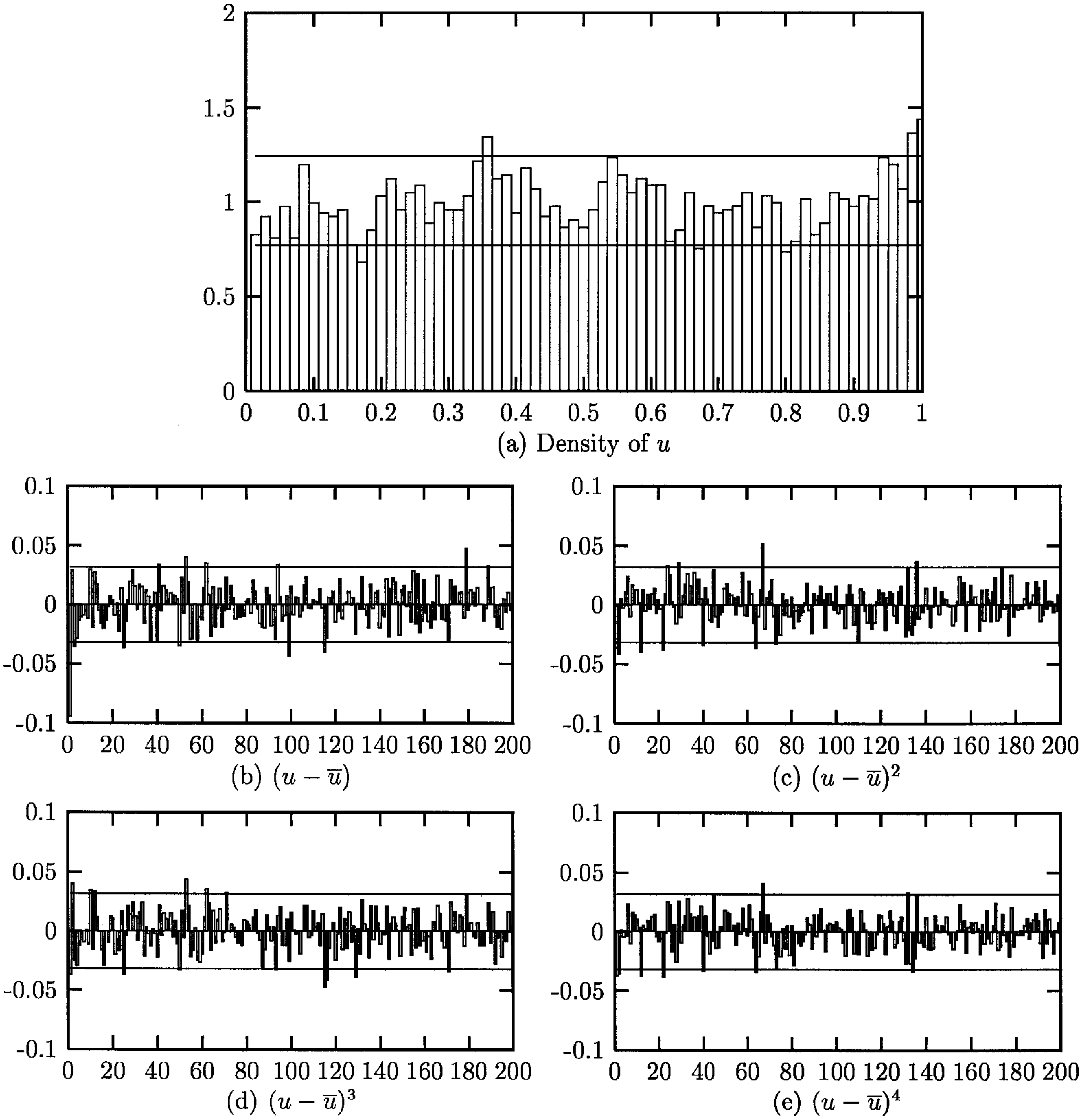
The horizontal lines superimposed on the histogram (a) are approximate 95% confidence intervals for the individual bin heights under the null that u_t is i.i.d U(0,1). The horizontal lines superimposed on the correlograms (b)-(e) are approximate 95% confidence intervals for the individual bin heights under the null that u_t is i.i.d U(0,1).

Figure 2.11: ARJI, Estimates of the density of u_t and autocorrelation functions of powers of u_t.



See the notes to Figure 2.10.

Figure 2.12: ARJI-h_t, Estimates of the density of u_t and autocorrelation functions of powers of u_t.



See the notes to Figure 2.10.

Chapter 3

A Correlated Bivariate Poisson Jump Model for Foreign Exchange

3.1 Introduction

Jump dynamics are an important element for modeling stock market movements and Chapter 2 provides evidence of these systematic jumps in financial markets. A natural extension is to develop a multivariate model that can be used to exploit information across equations and most importantly that can supply a reliable learning tool for investigating the comovement between financial asset returns.

The potential gain in the efficiency of estimation can be substantial depending on the underlying information structure. A multivariate model captures additional information contained in the covariance between random normal error terms as well as that contained in the covariance between different jumps experienced by financial assets. Incorporating the relationship between jump dynamics may provide significant improvements in efficiency of estimators.

In addition, modeling the similarity of jumps in different return series can enrich our understanding of the dynamic relationship between returns. It is natural to think that since two return series experience irregular jumps or crashes at irregular intervals, the crashes could be classified into two categories: (i) independent jumps -

crashes that affect one series specifically and not the others; and (ii) correlated jumps - crashes that impact all series simultaneously. In a financial market, there are news events that induce changes in one company alone. For example, a profit warning released by a firm will usually affect the value of its own stock alone. There are other news events that could shock the entire market such as reports on macroeconomic indicators.

A similar argument can be made for currency markets. Domestic news related to the local economy is one of many examples of news that may affect the value of one particular currency, whereas news such as that of a global financial crisis could easily transmit turbulence into several different currencies. In a global currency market, sudden and abnormal changes in any given currency can be classified as either independent jumps or correlated jumps.

This chapter develops a new bivariate jump model to study jump dynamics in foreign exchange returns. The model extends a multivariate GARCH parameterization to include a bivariate correlated jump process. The conditional covariance matrix has the Baba, Engle, Kraft, and Kroner (1989) structure, while the bivariate jumps are governed by a Correlated Bivariate Poisson (CBP) function.

There are several advantages of using this CBP-GARCH model. Firstly, it mixes smooth volatility movement with abrupt changes in returns. The incorporation of jumps provides one possible solution to account for unconditional leptokurtosis. Secondly, it allows one to identify two types of systematic jumps: jumps specific to one currency and jumps that affect both currencies simultaneously. Thirdly, it allows for the frequency of jumps to change over time, depending on market conditions. Fourthly, the interrelationship between currencies in such a model are driven by two distinct sources: normal random noises and systematic correlated jumps. Jump dynamics may provide a better understanding of the comovement between currencies, which has important implications for risk management and hedging such as deriving the optimal hedging ratio (Baillie and Myers 1991).

We apply this model to eleven years of data on daily spot exchange rates of the Canadian Dollar (CD) and Japanese Yen (JY) against the US dollar. We find systematic independent as well as correlated jumps, with significant jump size, in both currencies. While the Canadian dollar on average experiences more positive jumps, leading to depreciation, the Japanese Yen encounters mostly negative jumps. In our modeling efforts, we also propose two generalizations for the jump frequency allowing it to be time-varying. The first approach relates the likelihood of jumps to the market conditions so that the arrival of independent jumps is determined by the currency's volatility, whereas the arrival of correlated jumps is jointly determined by the volatilities in both currencies. The second approach allows the jump frequency to follow an approximate ARMA process driven by the unexpected number of jumps last period.

The chapter is organized as follow: Section 3.2 reviews the development of the literature. Section 3.3 describes the Correlated Bivariate Poisson (CBP) jump model. Section 3.4 provides a simple data description with summary statistics. Section 3.5 applies the CBP-GARCH model to the foreign exchange rates in our data set. Section 3.6 offers conclusions.

3.2 Literature Review

Empirical research has so far failed to find a predictable component in exchange rates using linear models. Purchasing power parity, flexible monetary models (Frenkel 1976), sticky price monetary models (Dornbusch 1976, Frenkel 1979), and vector autoregressive models often fail to outperform a simple random walk model. The results are documented in Meese and Rogoff (1983) and basically suggest that exchange rates are unpredictable and the behavior of foreign exchange markets is consistent with the efficient market hypothesis (Fama 1970).

The failure of linear exchange rate models has led to a shift of focus towards non-linear modeling. The Autoregressive Heteroskedasticity (ARCH) model, proposed in the seminal paper by Engle (1982) and later generalized (GARCH) by Bollerslev (1986), has been most influential. This parsimonious structure implies serial correlation in the second moment and volatility clustering, suggesting that periods of high (low) volatility are likely to be followed by periods of high (low) volatility. The univariate GARCH model has been used extensively in the foreign exchange literature (Domowitz and Hakkio 1985, Hsieh 1989a, Engle and Bollerslev 1986, McCurdy and Morgan 1988).

Multivariate GARCH models (Bollerslev 1990, Baba, Engle, Kraft, and Kroner 1989, Bollerslev, Engle, and Wooldridge 1988, Diebold and Nerlove 1989, Engle, Ng, and Rothschild 1990) emerged as a natural extension of the univariate model. The motivations behind the multivariate generalization are possible volatility spillover effects and a quest for an understanding of how one market might influence another. For example, if all currencies being studied are expressed in terms of a common denomination (U.S. dollar), any shock to the U.S. market may easily be transmitted to all currencies, producing similar GARCH effects.

Although multivariate GARCH models prove to be adequate in terms of accounting for heteroskedasticity, these models do not fully capture another stylized fact: leptokurtosis in the unconditional distribution, often observed in financial data. Fat tails in foreign exchange rates are documented in many studies including Burt, Kaen, and Booth (1977), Westerfield (1977), Rogalski and Vinso (1978), and Friedman and Vandersteel (1982). Many solutions have been proposed in the literature. For example, the normal density can be replaced by a fat tail distribution such as the Student t distribution or Power Exponential distribution. Other alternatives include the Poisson jump model of Press (1967) which introduces an independent jump process with the arrival of jumps governed by a Poisson distribution. This model has been applied successfully to daily exchange rates by Akgiray and Booth (1988), Tucker and Pond (1988) and Hsieh (1989b). This approach is attractive as more can be learned from modeling leptokurtosis as a systematic pattern than by simply utilizing a fat tail distribution. Although jumps are unobservable, an expost filter can always be constructed to infer the probability of jumps.

The presence of jumps can be explained either by news content entering the market or, more interestingly by market microstructure - order flow recently proposed by Evans and Lyons (2001). The former implies that market participants may react to certain types of unanticipated news systematically over time. Modeling these pattern is no easy task, the Poisson distribution provides a simple entry point which has proven to be useful in empirical studies. The market microstructure approach (Evans and Lyons 2001) relies on portfolio shifts not being common knowledge. Dealers observe interdealer order flow to learn about these shift. As the market gradually aggregates these pieces of information, the transactions between dealers and the non-dealer public may create a series of jumps in the exchange rates that they are trading.

This chapter proposes a new bivariate jump model to study jump dynamics in foreign exchange returns. A multivariate GARCH parameterization in the BEKK form augmented with a Correlated Bivariate Poisson (CBP) function is developed to allow for a bivariate correlated jump process. This CBP function provides a bivariate discrete counting process which has been used to solve problems in various context such as the relationship between voluntary and involuntary job changes (Jung and Winkelmann 1993), and the entry and exit decisions of firms (Mayer and Chappell 1992). Modeling of pairs of discrete dependent variables using the CBP function is also discussed in Gourieroux, Monfort, and Trognon (1984).

3.3 Model

The model is a combination of the GARCH model (Bollerslev 1986) and the Poisson Correlation function (M'Kendrick 1926, Campbell 1934). Given an information set, $\Phi_t = (R_t, ..., R_1)$, the Correlated Bivariate Poisson (CBP-GARCH) jump model is defined as follow:

$$R_t = \mu + \phi R_{t-1} + \epsilon_t + J_t \tag{3.1}$$

where R_t is a 2×1 vector of returns consisting of a constant mean μ (2×1), a random disturbance ϵ_t (2×1), and a jump component J_t (2×1). The random disturbance follows a bivariate normal distribution with zero mean and variance covariance matrix $\tilde{H}_t$.

Within any single time period t, a currency may experience "n" number of jumps, where "n" depends on the news content entering the market. The jump component therefore is constructed as a sum of a series of random variables Y_i:

$$\sum_{i=1}^{n} Y_i = Y_1 + Y_2 + Y_3 + ... + Y_n \tag{3.2}$$

Each of these random variables can be interpreted as a jump size which is governed by a normal distribution with constant mean θ and constant variance δ. We assume that these mean and variance parameters remain the same across time, but differ across currencies. In other words, the jump sizes for the two currencies can be characterized as

$$Y_{1t,i} \sim N(\theta_1, \delta_1^2) \quad \text{and} \quad Y_{2t,j} \sim N(\theta_2, \delta_2^2). \tag{3.3}$$

The jump component enters the mean equation with an expected value of zero which is achieved by subtracting the expected values from the series of random jumps.

In a bivariate framework, the jump component is defined as

$$J_t = \begin{bmatrix} \sum_{i=1}^{n_{1t}} Y_{1t,i} - E_{t-1}(\sum_{i=1}^{n_{1t}} Y_{1t,i}) \\ \sum_{j=1}^{n_{2t}} Y_{2t,j} - E_{t-1}(\sum_{j=1}^{n_{2t}} Y_{2t,j}) \end{bmatrix}. \tag{3.4}$$

J_t has a bivariate normal distribution with zero mean and variance covariance matrix Δ_t. The disturbance ϵ_t and the jump components are assumed to be independent.

3.3.1 The Poisson Correlation Function

Two discrete counting variables n_{1t} and n_{2t} control the arrival of jumps and they are constructed via three independent Poisson variables namely n_{1t}^*, n_{2t}^*, and n_{3t}^*. Each one of these variables has a probability density function given by

$$P(n_{it}^* = j|\Phi_{t-1}) = \frac{e^{-\lambda_i}\lambda_i^j}{j!}. \tag{3.5}$$

The expected value and variance of n_{it}^* are each equal to λ_i, which is also referred to as the expected number of jumps or the jump intensity.

The correlated jump intensity counters (M'Kendrick 1926, Campbell 1934) are defined as

$$n_{1t} = n_{1t}^* + n_{3t}^* \quad \text{and} \quad n_{2t} = n_{2t}^* + n_{3t}^*. \tag{3.6}$$

By construction, each of these counting variables, n_{it}, is capable of generating independent jumps as well as correlated jumps. The independent jumps are initiated by n_{1t}^* and n_{2t}^* in time period t. The correlated jumps are produced by the additional Poisson variable n_{3t}^* which contributes jumps to both series.

The joint probability density function for the three independent Poisson variables is

$$P(n_{1t}^* = i, n_{2t}^* = j, n_{3t}^* = k) = \frac{e^{-\lambda_1}\lambda_1^i}{i!}\frac{e^{-\lambda_2}\lambda_2^j}{j!}\frac{e^{-\lambda_3}\lambda_3^k}{k!}. \tag{3.7}$$

Using the change of variables method and integrating out n_{3t}^* yields the joint probability density for n_{1t} and n_{2t} as:

$$P(n_{1t} = i, n_{2t} = j|\Phi_{t-1}) = \sum_{k=0}^{min(i,j)} \mathrm{e}^{-(\lambda_1+\lambda_2+\lambda_3)} \frac{\lambda_1^{i-k}\lambda_2^{j-k}\lambda_3^k}{(i-k)!(j-k)!k!}. \tag{3.8}$$

This Poisson Correlation function has been studied extensively. Teicher (1954), and Hamdan and Al-Bayyati (1969) derive the Poisson Correlation function as the limit of a bivariate binomial distribution. Martiz (1952) derives the function using the factorial moment generating function of a bivariate Bernoulli distribution. Marshall and Olkin (1985) provide a detailed description of its properties. Maximum likelihood estimation is discussed in Holgate (1964).

The conditional expectation is given by

$$E(n_{1t}|n_{2t} = j) = \lambda_1 + j\frac{\lambda_3}{\lambda_2 + \lambda_3} \tag{3.9}$$

and the expected number of jumps is equal to

$$E(n_{it}) = \lambda_i + \lambda_3. \tag{3.10}$$

Since both n_{1t} and n_{2t} are monotone functions of independent Poisson random variables, their covariance is λ_3 and their correlation is always positive in the form of

$$\mathrm{Corr}(n_{1t}, n_{2t}) = \frac{\lambda_3}{\sqrt{(\lambda_1 + \lambda_3)(\lambda_2 + \lambda_3)}}. \tag{3.11}$$

The assumption of positive correlation can be relaxed by using other bivariate distributions. However, in this model, there is built-in flexibility in terms of jumps. Although the model assumes that the number of jumps in the two series are positively related, the mean of the jump sizes can be different. This implies that it is possible for both series to experience a jump at the same time, but that one jump may have a positive effect on one series and a negative effect on the other.

Note that the bivariate Poisson function simplifies to two independent Poisson process when λ_3=0. When either one of the independent jump intensities λ_1 or λ_2 is equal to zero, the distribution is called semi-Poisson. In the situation of $\lambda_1 = 0$, positive probability is only assigned to the situation where $n_{1t} \leq n_{2t}$. In the case of $\lambda_1 = \lambda_2 = 0$, the only possible values that the number of jumps can take is $n_{1t} = n_{2t}$.

A drawback of using the Poisson Correlated function is that the correlation is limited to the range:

$$0 \leq \mathrm{Corr}(n_{1t}, n_{2t}) \leq \min \left[\left(\frac{\lambda_1 + \lambda_3}{\lambda_2 + \lambda_3} \right)^{1/2}, \left(\frac{\lambda_2 + \lambda_3}{\lambda_1 + \lambda_3} \right)^{1/2} \right]. \tag{3.12}$$

Given a constant value for each jump intensity parameter λ_i, the correlation between n_{1t} and n_{2t} will have a theoretically imposed upper bound. This assumption can be relaxed. Aitchison and Ho (1989) assume the jump intensity parameter itself has a log normal distribution. In a similar fashion, Edwards and Gurland (1961) propose a compound correlated Poisson distribution arising from compounding a CBP function with Gamma distribution. Although these multivariate mixtures of independent Poisson distributions allow for a rich covariance structure for the counting variables, practical applications for regression analysis is limited due to the difficulty of estimation and inference.[1]

3.3.2 Time Varying Jump Intensity

To relax the assumption of constant correlation and the implicitly imposed upper bound, we propose two simple parametric structures for each of the λ_i, so that the likelihood of jumps may change over time. The first approach links the variations of jump intensity to the market conditions approximated by last period's volatility. The likelihood of jumps rises as the market becomes more volatile. The second approach

[1]See Holgate (1964) and Aitchison and Ho (1989) for more discussion.

allows the jump intensity to follow a parsimonious ARMA structure governed by the unexpected number of jumps last period.

3.3.2.1 Conditioning on Last Period's Volatility

The time varying jump intensities (CBP-GARCH-R_{t-1}^2) conditioning on last period's volatility are defined as

$$\begin{aligned} \lambda_{1t} &= \lambda_1 + \eta_1^2 r_{1t-1}^2, & (3.13) \\ \lambda_{2t} &= \lambda_2 + \eta_2^2 r_{2t-1}^2, & (3.14) \\ \lambda_{3t} &= \lambda_3 + \eta_3^2 r_{1t-1}^2 + \eta_4^2 r_{2t-1}^2, & (3.15) \end{aligned}$$

where r_{it-1} is the rate of return for currency i at time t-1. The jump intensities are assumed to be related to market conditions which are reflected in r_{it-1}^2 as an approximation of last period's volatility.[2] Similarly, the covariance λ_{3t} is governed by the variations in last period volatilities from both series. This parametric structure not only introduces additional jump dynamics to the model, but also allows for a time-varying correlation between the counting variables n_{1t} and n_{2t}. Since λ_{it} is realized at time t-1, one step ahead forecasts of the average number of jumps is also possible.

A drawback of using this simple structure is the implicitly imposed deterministic relationship between the intensities. Since all three intensity parameters are related to the same set of variables (previous period volatilities from both currencies), the

[2]The jump intensities are positively related to the volatilities. This assumption can be relaxed by using logarithms such as

$$\begin{aligned} \lambda_{1t} &= \exp\left[\lambda_1 + \eta_1 \ln r_{1t-1}^2\right] & (3.16) \\ \lambda_{2t} &= \exp\left[\lambda_2 + \eta_2 \ln r_{2t-1}^2\right] & (3.17) \\ \lambda_{3t} &= \exp\left[\lambda_3 + \eta_3 \ln r_{1t-1}^2 + \eta_4 \ln r_{2t-1}^2\right] & (3.18) \end{aligned}$$

This specification allows for both positive and negative correlation between the jump intensities and past volatilities.

comovement between the series is restricted to a linear structure:

$$\lambda_{3t} \;=\; \lambda_3 + \eta_3^2\left(\frac{\lambda_{1t}-\lambda_1}{\eta_1^2}\right) + \eta_4^2\left(\frac{\lambda_{2t}-\lambda_2}{\eta_2^2}\right) \tag{3.19}$$

which will hold in any time period. A similar problem arises as we replace the past volatility with any other exogenous variables maintaining the assumption that the variables affecting the likelihood of independent jumps should also simultaneously affect the likelihood of correlated jumps. A possible solution is to allow for a counter-specific stochastic component entering the jump intensity similar to the stochastic volatility model. However, this structure will certainly complicate estimation and inference procedures, quite often requiring simulation methods. A more simplified model is desirable and the autoregressive conditional jump intensity to be developed in next subsection is one such model.

3.3.2.2 Autoregressive Conditional Jump Intensity (ARJI)

The Autoregressive Conditional Jump Intensity (ARJI) model was first introduced in Chapter 2 and it has proven to work well with stock market series in terms of volatility forecasting. A practical challenge is to extend this model into a multivariate framework. We will discuss a bivariate extension in this chapter. However, a multivariate model can be derived without any difficulty following the same steps as discussed below.

Let $\lambda_t \equiv E[n_t^*|\Phi_{t-1}]$ be the conditional expectation of the counting process which is assumed to follow,

$$\lambda_t = \lambda_0 + \sum_{i=1}^{r} \zeta_i \lambda_{t-i} + \sum_{i=1}^{s} \gamma_i \xi_{t-i}. \tag{3.20}$$

The conditional jump intensity at time t is related to r past lags of the conditional jump intensity plus lags of ξ_t. ξ_{t-i} represents the innovation to λ_{t-i} as measured *ex post* by the econometrician. This *shock*, or jump intensity residual, is calculated as follows,

$$\begin{aligned} \xi_{t-i} &\equiv E[n^*_{t-i}|\Phi_{t-i}] - \lambda_{t-i} \\ &= \sum_{j=0}^{\infty} jP(n^*_{t-i} = j|\Phi_{t-i}) - \lambda_{t-i} \end{aligned} \tag{3.21}$$

The first term on the right hand side of Equation 3.21 is our inference on the average number of jumps at time $t-i$ based on time $t-i$ information, while the second term in (3.21) is our expectation of the number of jumps using information at time $t-i-1$. Therefore, ξ_{t-i} represents the unpredictable component affecting our inference about the conditional mean of the counting process n^*_{t-i}.

Let $f(R_t|n^*_t = j, \Phi_{t-1})$ denote the conditional density of returns given j jumps occur, and the information set Φ_{t-1}; where f represents any density function. Having observed R_t and using Bayes' rule we can infer ex post the probability that a jump of size j occurred at time t, with the filter defined as,

$$P(n^*_t = j|\Phi_t) = \frac{f(R_t|n^*_t = j, \Phi_{t-1})P(n^*_t = j|\Phi_{t-1})}{P(R_t|\Phi_{t-1})}, \quad j = 0, 1, 2, \ldots \tag{3.22}$$

where $P(n^*_t = j|\Phi_{t-1})$ is defined in Equation 3.5. The key innovation of this ARJI model in a bivariate framework is to trace the path of unexpected jump intensities through the independent jump counters n^*_t instead of the the counters n_t after the variable transformation. The conditional density in Equation 3.22 can be easily derived by integrating out the nuisance variables. For example, the density of returns given j correlated jumps can be obtained as

$$f(R_t|n^*_{3t} = k, \Phi_{t-1}) = \sum_i \sum_j f(R_t|n^*_{1t} = i, n^*_{2t} = j, n^*_{3t} = k, \Phi_{t-1}). \tag{3.23}$$

The filter in Equation (3.22) is an important component of our model of time-varying jump dynamics, since it enters Equation (3.21), but it also can be constructed and used for inference purposes. For example, the probability that at least one jump occurred in currency 1 could be assessed using $1 - P(n^*_{1t} = 0|\Phi_t)$. The filter may be particularly useful in revealing misspecification in the simpler constant intensity specification.

3.3.3 The CBP-BEKK Variance Covariance Matrix

Combining the GARCH model with the CBP function, the probability density function for R_t given i jumps in currency 1 and j jumps in currency 2 is defined by

$$f(R_t|n_{1t}=i, n_{2t}=j, \Phi_{t-1}) = \frac{1}{2\pi^{N/2}}|H_{ij,t}|^{-1/2} exp\left[-u'_{ij,t}H_{ij,t}^{-1}u_{ij,t}\right], \tag{3.24}$$

where $u_{ij,t}$ is the usual error term with the jump component $J_{ij,t}$ representing the effect of i and j jumps:

$$u_{ij,t} = R_t - \mu - J_{ij,t} = \begin{bmatrix} r_{1t} - \mu_1 - i\theta_1 + (\lambda_{1t} + \lambda_{3t})\theta_1 \\ r_{2t} - \mu_2 - j\theta_2 + (\lambda_{2t} + \lambda_{3t})\theta_2 \end{bmatrix}. \tag{3.25}$$

$H_{ij,t}$ is the variance covariance matrix of the returns given i jumps in currency 1 and j jumps in currency 2. Under the assumption that the normal disturbance, ϵ_t, is independent of the jump component, $H_{ij,t}$ can be separated into two parts: the variance covariance matrix for the normal random disturbance $\tilde{H}_t$ and the variance covariance matrix for the jump component $\Delta_{ij,t}$.

The variance covariance matrix $\tilde{H}_t$ for the normal disturbance is assumed to have the bivariate BEKK form (Baba, Engle, Kraft, and Kroner 1989) which is defined by

$$\tilde{H}_t = C'C + A'\epsilon_{t-1}\epsilon'_{t-1}A + B'\tilde{H}_{t-1}B \tag{3.26}$$

where C is an upper triangular matrix, and A and B are 2×2 parameter matrices. The positive definiteness of the variance covariance matrix is ensured by the quadratic form. This structure allows for a very flexible relationship between the second moments and their past values. Individual elements of $\tilde{H}_t$ can be written as

$$h_{1,t} = c_1 + a_1^2u_{1,t-1}^2 + 2a_1a_{21}u_{1,t-1}u_{2,t-1} + a_{21}^2u_{2,t-1}^2$$

$$
\begin{aligned}
& +b_1^2 h_{1,t-1} + 2b_1 b_{21} h_{12,t-1} + b_{21}^2 h_{2,t-1}, && (3.27) \\
h_{12,t} \;=\; & c_{12} + a_1 a_{12} u_{1,t-1}^2 + (a_1 a_2 + a_{12} a_{21}) u_{1,t-1} u_{2,t-1} + a_2 a_{21} u_{2,t-1}^2 && \\
& +b_1 b_{12} h_{1,t-1} + (b_1 b_2 + b_{12} b_{21}) h_{12,t-1} + b_2 b_{21} h_{2,t-1}, && (3.28) \\
h_{2,t} \;=\; & c_2 + a_2^2 u_{2,t-1}^2 + 2a_2 a_{12} u_{1,t-1} u_{2,t-1} + a_{12}^2 u_{1,t-1}^2 && \\
& +b_2^2 h_{2,t-1} + 2b_2 b_{12} h_{12,t-1} + b_{12}^2 h_{1,t-1}. && (3.29)
\end{aligned}
$$

Note that both variance and covariance terms have similar structures. They are functions of the past variances, prediction errors, past covariances, and cross products of the prediction errors.

This BEKK formulation provides a rich class of variance covariance structures for the bivariate model and it also nests Bollerslev, Engle, and Wooldridge (1988)'s diagonal vech parameterization as a special case when A and B are diagonal matrices (or equivalently $a_{12} = a_{21} = b_{12} = b_{21} = 0$).

The variance covariance matrix for the jump component $\Delta_{ij,t}$ is derived from the assumption that the correlation between the jump sizes is constant across contemporaneous equations and zero across time:

$$
Corr(Y_{1t}, Y_{2t}) = \rho_{12} \quad \text{and} \quad Corr(Y_{1t}, Y_{2s}) = 0 \quad t \neq s \tag{3.30}
$$

Therefore, conditional on i and j jumps, the variances of the jump components $\sum_i Y_{1t}$ and $\sum_j Y_{2t}$ are $i\delta_1^2$ and $j\delta_2^2$, respectively. The covariance will be $\rho_{12}\sqrt{ij}\delta_1\delta_2$ which completes the specification of the variance covariance matrix for the jump component as

$$
\Delta_{ij,t} \;=\; \begin{bmatrix} i\delta_1^2 & \rho_{12}\sqrt{ij}\delta_1\delta_2 \\ \rho_{12}\sqrt{ij}\delta_1\delta_2 & j\delta_2^2 \end{bmatrix}. \tag{3.31}
$$

The variance covariance matrix for the CBP-GARCH model $H_{ij,t}$ will always be positive definite as long as $\tilde{H}_t$ is positive definite. By construction, the variance covariance matrix for the jump component $\Delta_{ij,t}$ is well defined given i and j jumps and therefore $H_{ij,t}$ as the sum of two positive definite matrices will also be positive definite.

Introducing correlated jumps in the bivariate model has important implications for the covariance between the currencies. Assuming that there are no jumps in the series, the covariance will be measured by the off-diagonal elements in the BEKK structure as $h_{12,t}$. The time-varying covariance will be solely determined by the characteristics of the normal disturbances. In the presence of jumps, the covariance between two currencies will be driven not only by the covariance between the normal disturbances, but also by the characteristics of the jumps.

In terms of portfolio diversification, the correlation between the jump sizes provides important information about the correlation between the returns. For example, a positive correlation, ρ, implies increasing number of jumps will cause movements in both currencies in the same direction, and vice versa. In other words, as the market becomes more volatile, the benefit of diversification diminishes because the *shocks* tend to drive the currencies to move in a similar fashion.

With simple algebraic manipulation, we can show that the first four moments of the returns for currency i are given by

$$\mathrm{E}(r_{it}|\Phi_{t-1}) = \mu, \tag{3.32}$$

$$\mathrm{Var}(r_{it}|\Phi_{t-1}) = h_{i,t} + (\lambda_{it} + \lambda_{3t})(\delta_i^2 + \theta_i^2), \tag{3.33}$$

$$\mathrm{Skew}(r_{it}|\Phi_{t-1}) = \frac{(\lambda_{it} + \lambda_{3t})(\theta_i^3 + 3\theta_i\delta_i^2)}{(h_{i,t} + (\lambda_{it} + \lambda_{3t})\delta_i^2 + (\lambda_{it} + \lambda_{3t})\theta_i^2)^{3/2}}, \tag{3.34}$$

$$\mathrm{Kurt}(r_{it}|\Phi_{t-1}) = 3 + \frac{(\lambda_{it} + \lambda_{3t})(\theta_i^4 + 6\theta_i^2\delta_i^2 + 3\delta_i^4)}{(h_{i,t} + (\lambda_{it} + \lambda_{3t})\delta_i^2 + (\lambda_{it} + \lambda_{3t})\theta_i^2)^2}. \tag{3.35}$$

The conditional variance is composed of the GARCH variance and the parameters characterizing the jumps. The conditional volatility is positively related with the jump intensity, the variance of the jump size, and the magnitude of the jump size. Increasing the number of jumps in the returns will result in higher volatilities. The exact nature of the change in volatility will depend on the size of these jumps. Either positive or negative jumps with significantly large jump size will cause the volatility to rise. If jump size is characterized by a large variance, jumps entering the market will tend to be of very different sizes creating higher volatilities in returns.

3.3.4 *Ex post* Filter and the Likelihood Function

Finally, to complete the specification, the conditional density of returns is defined by

$$P(R_t|\Phi_{t-1}) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} f(R_t|n_{1t}=i, n_{2t}=j, \Phi_{t-1})P(n_{1t}=i, n_{2t}=j|\Phi_{t-1}). \tag{3.36}$$

Although jumps are unobservable, an *ex post* filter can be constructed as

$$\begin{aligned} P(n_{1t}=i, n_{2t}=j|\Phi_t) &= \frac{f(R_t|n_{1t}=i, n_{2t}=j, \Phi_{t-1})}{P(R_t|\Phi_{t-1})} \times \\ &\quad P(n_{1t}=i, n_{2t}=j|\Phi_{t-1}) \end{aligned} \tag{3.37}$$

to identify jumps in the series. The log likelihood function is simply the sum of the log conditional densities:

$$\ln L = \sum_{t=1}^{N} \ln P(R_t|\Phi_{t-1}). \tag{3.38}$$

Since the information matrix is not block diagonal, evaluation of the full likelihood is required.

For the case of the ARJI model, we can simply redefine the probability density function of R_t given i independent jumps in currency 1, j independent jumps in currency 2, and k correlated jumps in both currencies as

$$\begin{aligned} f(R_t|n^*_{1t}=i, n^*_{2t}=j, n^*_{3t}=k, \Phi_{t-1}) &= \frac{1}{2\pi^{N/2}}|H_{ijk,t}|^{-1/2} \times \\ &\quad exp\left[-u'_{ijk,t}H^{-1}_{ijk,t}u_{ijk,t}\right], \end{aligned} \tag{3.39}$$

where $u_{ijk,t}$ is given by

$$u_{ijk,t} = R_t - \mu - J_{ijk,t} = \begin{bmatrix} r_{1t} - \mu_1 - (i+k)\theta_1 + (\lambda_{1t} + \lambda_{3t})\theta_1 \\ r_{2t} - \mu_2 - (j+k)\theta_2 + (\lambda_{2t} + \lambda_{3t})\theta_2 \end{bmatrix} \tag{3.40}$$

and the variance covariance matrix for the jump component is

$$\Delta_{ijk,t} = \begin{bmatrix} (i+k)\delta_1^2 & \rho_{12}\sqrt{(i+k)(j+k)}\delta_1\delta_2 \\ \rho_{12}\sqrt{(i+k)(j+k)}\delta_1\delta_2 & (j+k)\delta_2^2 \end{bmatrix}. \tag{3.41}$$

The likelihood function will be the products of this conditional density and the independent Poisson function (3.5).

To estimate the CBP-GARCH model, a truncation point must be selected for the probability function in Equation (3.36). We choose a truncation point which is sufficiently large so that the likelihood function and parameter estimates stabilize to a set of converged values.

3.4 Data

We use daily spot exchange rates for the Canadian Dollar (CD) and Japanese Yen (JY) relative to the U.S. Dollar from January 2, 1990 to December 29, 2000. Table 3.1 provides summary statistics for the rates of return expressed as one hundred times the first differenced logarithms.

The data set covers episodes of many currency crises including the near-breakdown of the European Exchange Rate Mechanism (1992-1993), the Latin American Tequila Crisis following Mexico's peso devaluation (1994-1995), the Asian Crisis (1997-1998), and the Russian Crisis (1998). During the Asian Financial Crisis, the Japanese Yen experienced only a moderate devaluation between July 1997 and January 1998. The two currencies were mostly unaffected by these episodes of turmoil. The Bank of Japan has intervened in some occasions. For example, the Bank of Japan intervened in the market to support yen on June 30, 1998 and to support U. S. dollar on April 3, 2000.

The summary statistics in Table 3.1 show that there is significant excess kurtosis in the CD and the modified Ljung Box (West and Cho 1995) statistics indicate serial correlation in both the first and second moments of the returns. Similar descriptive statistics are found for the JY. Serial correlation exists in the squared returns. The excess kurtosis coefficient is slightly larger for the JY as compared to that for the CD. The Augmented Dickey-Fuller test shows that the two return series are stationary.

Figure 3.1 and 3.2 present plots of spot exchange rates and returns for the two currencies. A general trend of depreciation of the Canadian dollar is apparent for the last decade, while the Japanese Yen has experienced both periods of appreciation and depreciation over the sample period. Volatility clustering is common in both return series and therefore suggests the plausibility of the GARCH structure. There is no close relationship between the two series as the simple correlation coefficient between the returns is -0.02.

3.5 Results

We will first present the results of Jorion's (1988) Poisson-GARCH model to get a preliminary idea of the jump structure. The bivariate BEKK model with no jumps is estimated next to examine the variance covariance structure. The results from the CBP-GARCH model will then be presented with a detailed discussion of the jump component. The extended model with time-varying jump intensities is reported at the end of this section.

3.5.1 Poisson Jumps and the BEKK Structure

The results for the Poisson-GARCH model are presented in Table 3.2. The first thing to note are the strong GARCH effects and the persistence of the conditional variance, with parameters $\alpha + \beta$=0.9752 for CD and 0.9817 for JY. The Poisson jump components are also very important in modeling these exchange rates as shown by the significance of both the jump intensity and size parameters. The estimated jump intensities are 0.2142 and 0.2056 for CD and JY, respectively. The jumps in CD are mostly positive with a small variance, whereas the jumps in JY have a negative mean and a relatively larger variance. We found it necessary to use an AR(1) to capture serial correlation in the conditional mean of exchange rate returns for all models: the corresponding coefficients are denoted as ϕ_i. Overall, these results indicate that although the (previously reported) simple correlation coefficient (-0.02) shows no indication of a relationship between the returns, both exchange rate return series exhibit systematic jumps. These jump structures may be different, however, any attempt to model the two series jointly must take into account these jump components.

The bivariate BEKK model reported in Table 3.3 shows a similar GARCH effect to those in Table 3.2, as the conditional variances are autocorrelated with significant parameters a_1 and b_1 for CD and a_2 and b_2 for JY. However, based on Equations (3.27)-(3.29), we find no relationship between the conditional variances and covariance as most of the off-diagonal parameters from the A and B matrices are small and insignificant. The insignificance of these parameters suggests that the conditional variance is not affected by the past conditional covariance and similarly the conditional covariance is not affected by the past conditional variance. The relationship between returns may be complex. However the strong GARCH effect can always be identified with either a univariate model or a bivariate BEKK model.

3.5.2 The Correlated Bivariate Poisson GARCH Model

The estimates of the CBP-GARCH model with constant jump intensities are presented in Table 3.4. The constants, μ_i, are insignificant for both series and again strong GARCH effects are present. In comparison with the BEKK model, the

GARCH structures are essentially the same. The estimated cross equation parameters are all insignificant. The parameters associated with the autoregressive conditional variance, b_1 and b_2, are of the same order of magnitude: with estimates of 0.9667 and 0.9730 for the BEKK model and 0.9577 and 0.9821 for the CBP-GARCH model. Similarly, the estimated parameters on the squared past prediction errors $\hat{a}_1$ and $\hat{a}_2$ are 0.2277 and 0.1946 for the BEKK model and 0.2364 and 0.1374 for the CBP-GARCH model. These results suggest that adding a jump component does not qualitatively or quantitatively affect the conditional variance structure of the random disturbance. The CBP-GARCH model does, however, use additional information (jump dynamics) already available in the data series to better understand the comovement between the returns.

Turning to the jump component, all parameters related to the jump dynamics are highly significant. The likelihood ratio test statistic for the presence of jumps is 408.74 with p-value equal to 0.0.[3] The mean θ and variance δ of the jump sizes have values closely resembling the results from the univariate Poisson-GARCH model. For CD, the mean and variance are 0.0515 and 0.3311 for the univariate case and 0.0526 and 0.3381 for the CBP-GARCH model. For JY, the mean and variance are -0.2669 and 0.9157 for the univariate case and -0.2784 and 0.9477 for the bivariate case. Note that even after introducing the correlated jumps, the two currencies still appear to have opposite jump sizes with CD experiencing mostly positive jumps in the last decade and JY encountering mostly negative jumps which are five times the size of the ones in CD on average.

The jump intensities provide the keys to dividing the jumps into independent and

[3]Testing for jumps is complicated by the lack of identification of the nuisance parameters θ_i and δ_i under the null hypothesis. Drost, Nijman, and Werker (1996) propose a kurtosis-based test and Khalaf, Saphores, and Bilodeau (2000) suggest a Monte Carlo test to find an exact p-value. However, the power of these tests is not clear.

correlated components. In the univariate Poisson GARCH model, the jump intensity (the average number of jumps) in CD and JY are reported as 0.2142 and 0.2056, respectively. The equivalent measure in the CBP-GARCH model is $\hat{\lambda}_1 + \hat{\lambda}_3$=0.2276 for CD and $\hat{\lambda}_2 + \hat{\lambda}_3$=0.2002 for JY. Since independent jumps are initiated by λ_1 and λ_2, the average numbers of independent jumps in CD and JY are 0.1415 and 0.1141, respectively. The average number of correlated jumps, λ_3, in both series is estimated as 0.0861, implying that almost half the jumps occurring in both series are correlated, while the other half are independent.

There is significant correlation in the arrival of jumps between the two currencies which is shown by Corr(n_{1t}, n_{2t})=0.4033. The joint probabilities of jumps are depicted in panel (A) of Figure 3.5. Given the set of estimated intensity parameters from the CBP-GARCH model, the joint probabilities of jumps are centered at the origin and the probability of having over two jumps in both currencies on the same day is very small. This is a reasonable result given the expected number of jumps on average are 0.2276 in CD and 0.2002 in JY which are not even close to one jump. Another possible explanation is that this result is due to the fact that we assume a constant correlation between the counting variables across time. More realistically, the probability of jumps may change depending on market conditions. This provides motivation for considering the possibility that generalizing to time-varying jump intensities may shed further light on the relationships of currrency returns.

Although the correlation between jump sizes ρ is not statistically significant, the LR statistic for the null hypothesis of no correlated jumps $H_0 : \rho = 0$ and $\lambda_3 = 0$ is 11.23 with p-value equal to 0.003. The specification tests shows no uncaptured serial correlations in the square and cross product of the standardized residuals. However, the Ljung Box statistic on the jump intensity residuals (Q_{ξ_i}) rejects the assumption of constant correlated jump intensity implying the need for time-varying jump intensity.

3.5.3 Time Varying Jump Intensity

We next generalize the jump intensity parameters to be time-varying and positively related to previous volatilities as specified in Equation (3.13)-(3.15). Table 3.5 presents the estimates of the CBP-GARCH-R_{t-1}^2 model with time-varying jump intensities. Once again note that the GARCH effect remains in the series and all the estimated parameters of the GARCH structure are similar to those reported in the CBP-GARCH model.

The jump components are significant and the jump intensity parameters $\hat{\lambda}_i$ remain in the same range as before, with differences of less than 0.01 between the results from the two models. The most interesting results relate to the time-varying parameters which are on average very large in size and highly significant. Take for example the case of CD, the effect of last period's volatility on jump intensity η_1^2 is twice the size of the constant term $\hat{\lambda}_1$=0.1482. In other words, the likelihood of having jumps is directly related to the market conditions, reflected by the change in volatilities.

Figure 3.3 graphs the jump intensity dynamics which govern the independent and correlated jumps. Both $\hat{\lambda}_1$ and $\hat{\lambda}_2$ exhibit high variations in the sample period, and the assumption of constant intensities is clearly invalid. The jump intensities for CD and JY vary around the range of 0.1 to 1.1 and 0.1 to 1.55, respectively. It is not uncommon to have more than 0.5 jumps in any of the two currencies in one single day. Surprisingly the correlated jump intensity $\hat{\lambda}_3$ seems to exhibit variations with higher magnitude compared to the other two independent jump intensities. Note that $\hat{\lambda}_3$ has values as low as 0.1145 and as high as 2.9641. The results imply that the correlated jump dynamics are very important in understanding the comovement between the two currencies.

The joint probabilities created by these time-varying jump intensities are depicted in panel (B) of Figure 3.5. We have chosen the values $\lambda_1 = 0.5699$, $\lambda_2 = 0.8696$ and

$\lambda_3 = 2.9641$ on October 8, 1998 to examine the effect of relaxing the assumption of constant intensities on the joint probabilities. The graph clearly shows that the joint probabilities are centered around three jumps in both currencies. In fact, if we look at a particular number, the probability of having 2 jumps in CD and 3 jumps in JY is 0.05. It is not uncommon to have 2 or 3 jumps jointly in both currencies depending on the date. This indicates that the correlated jump dynamics seem to be a source of the changing relationship between the two currencies.

The same conclusion can be made by looking at the correlation between n_{1t} and n_{2t} in Panel (A) of Figure 3.6. The correlation varies from 0.3655 to 0.8489 and the assumption of constant correlation is clearly rejected. Diagnostic statistics show no serial correlation in standardized residuals or squared standardized residuals. The likelihood ratio test statistic of no time-varying jump dynamics yields a value of 25.86 which leads to rejection of the null hypothesis using the critical value from the χ^2 distribution with four degrees of freedom.

3.5.4 Autoregressive Conditional Jump Intensity (ARJI)

The results for the CBP-GARCH-ARJI model are reported in Table 3.6. A significant improvement of the likelihood compared to the CBP-GARCH-R^2 model is apparent and the likelihood ratio test for the null of constant intensity is again rejected. Allowing for time-varying jump intensity again results in a set of standardized residuals that are free of serial correlation.

A careful examination of the jump intensity estimates indicates consistency with the previous results. Recall that the correlated jump intensity residuals demonstrated serial correlation in the CBP-GARCH model (Table 3.4). The coefficients corresponding to the autoregressive structure are highly significant for the correlated jump intensity and not for either one of the independent jump intensities. The lag intensity coefficients ζ_i for the independent jumps are close to zero, whereas the same coefficient for the correlated jump intensity is significant and close to unity. High persistence in the correlated jump intensity implies any simultaneous *shocks* to both currencies will have a long lasting effect on the likelihood of having more disturbances in the future.

The time-varying CBP-GARCH-ARJI jump intensities are depicted in Figure 3.4 and the similarity between this set of plots as compared to the ones from the CBP-GARCH-R_{t-1}^2 model immediately appears. The correlated intensity $\hat{\lambda}_3$ lies in the range of 0.0344 to 3.1985 while the same measure from the the CBP-GARCH-R_{t-1}^2 model yields a range from 0.0756 to 2.9641. Although the three intensities do exhibit similar relative size in the two models, the intensities in the ARJI model seem to experience higher variations over the time period. A noticeable difference is the fact that each intensity seems to follow an unique path in the ARJI model, while the intensities driven by a common set of previous period volatilities show similar movements especially for the independent jumps governing Japanese Yen $\hat{\lambda}_2$ and the correlated jumps $\hat{\lambda}_3$.

As an illustration, consider the weakness of the U.S. stock market in September 1998. This led to a series of losses in the value of the U.S. dollar *vis-à-vis* other currencies. The Japanese Yen posted a surprising record gain of 5.71% in one single day on October 7. It was the biggest one day jump in 25 years and it erased nearly a year's worth of losses against the dollar. Similarly, the Canadian dollar recorded a 1.4% appreciation on the same day. This series of events will naturally lead to an increase in the likelihood of correlated jumps as opposed to the independent jumps.

A definite advantage of the CBP-GARCH-ARJI model is demonstrated by the joint probability function in Panel (C) of Figure 3.5, where the probability of having the same number of jumps in both currencies is clearly higher than the other combinations on October 8, 1998. In comparison, the linearly related intensities estimated from the CBP-GARCH-R_{t-1}^2 model, although they show a similar pattern force the combinations around the same number of jumps to increase together. Assuming the base currency U.S. dollar is experiencing some sort of internal *shock*, this will lead to correlated jumps in both currencies. Naturally, in such a situation, the probabilities of having correlated jumps increases, while the probabilities of having independent jumps remain the same. This can not be captured with a linear relationship between the jump intensities as modeled in the CBP-GARCH-R_{t-1}^2 model. The autoregressive conditional jump intensity supplies additional flexibility to modeling joint probability of jumps.

Comparison of the two plots in Figure 3.6 provides additional evidence of the superiority of the autoregressive structure as a modeling tool for time-varying intensity. The correlation between n_{1t} and n_{2t} in Panel (B) shows a substantially higher variation than the one in Panel (A). In addition, the correlation from the ARJI model covers almost the entire possible range of 0.0 to 1.0, whereas the same series from the R_{t-1}^2 model seems to be bounded within a much smaller interval of 0.5 between 0.35 to 0.85. It is likely that the bounds are created by the implicitly imposed upper bound from the Correlated Bivariate Poisson function as discussed in Section 3.3 along with the linear relationship between intensities as in Equation (3.19). The autoregressive structure effectively removes the theoretical upper bound imposed by the Poisson Correlation function.

The correlation between currencies in a state of high volatility is always an interesting question in empirical finance. The relationships of the jump intensities and the correlation between currencies are depicted as scatterplots in Figure 3.7. There are no obvious patterns for the independent jumps in the first two plots. As the number of independent jumps goes up, the correlation between currencies may change in different directions. Although the correlation coefficient, ρ, between jump sizes is negative and insignificant, a clear pattern can be found as the size of the correlated jump intensity becomes large. From the bottom scatterplot in Figure 3.7, we see that the correlation between currencies increases as the number of correlated jumps exceeds 1.5. The correlation between returns in a state of high volatility rises to the positive region once the number of *shocks* exceeds a certain threshold, creating risks that are difficult to diversify.

As a final specification check, root mean square errors (RMSE) and mean absolute errors (MAE) are examined to evaluate the in-sample and out-of-sample forecasting performance of conditional volatilities. The out-of-sample data cover the period 2 January 2001 to 31 August 2001, for a total of 170 observations. The actual standard deviation is approximated by the absolute deviation from the sample mean $|r_t - \bar{r}|$, the same approach as taken in So, Lam, and Li (1999). The results are presented in Table 3.7. A marginal improvement can be found in any of the jump models as compared to the plain BEKK-GARCH result. It is difficult however to distinguish the performance between the time-varying intensity models. Although, the CBP-GARCH-ARJI model is clearly superior to the other models according to the MAE measure. The improvements of volatility forecasts in RMSE and MAE provide evidence that the bivariate jump models do not overfit the data set.

In summary, the BEKK structure is adequate in modeling the conditional variance covariance structure between CD and JY. However, a jump component must be added to the model in order to fully capture the dynamics in the mean equation. With the Poisson Correlation function, we are able to identify independent as well as correlated jumps in the two currencies. A further generalization also discovers that the correlated jump dynamics may evolve over time which may serve as a good indicator for the future movement of the currencies.

3.6 Conclusion

In this chapter, we propose a bivariate GARCH model with jumps, with an application to the foreign exchange market. The BEKK structure is adopted for the conditional variance covariance matrix and the jump component is governed by a Poisson Correlation function. The model is applied to eleven years of data on daily spot exchange rates for the Canadian Dollar and Japanese Yen against the U.S. dollar.

The model provides several improvements over existing models. First, the CBP-GARCH model combines the popular multivariate GARCH model with a jump component so that it can capture smooth volatility movements as well as abrupt changes in the rates of return. Second, using the Poisson Correlation function to govern the jump component, the model is able to generate correlated jumps in both series in addition to the independent jumps. Third, we generalize the model to have time-varying jump intensities controlling the arrival of jumps which helps our understanding of the relationship between the jump dynamics and volatilities. Finally, allowing time-varying jump intensities also relaxes the assumption of constant and bounded jump correlation between currencies. This CBP-GARCH model has potential for use of asset pricing, modeling risk premia in foreign currency futures, and the modeling of optimal commodity hedge ratios.

Our results illustrate the empirical properties of two foreign exchange rates: the Canadian dollar and Japanese Yen. There are significant independent as well as correlated jumps in both series. The time-varying jump frequencies may provide important information as to how the correlation between the two currencies may evolve over time.

Table 3.1: Summary Statistics (1990 - 2000)

Canadian Dollar

	Mean	Standard Deviation	Skewness	Excess Kurtosis	Q(12)
R_t	0.0092	0.2985	-0.0044	5.4632	30.5
	(0.0056)	(0.0059)	(0.1294)	(0.4040)	[0.002]
$\|R_t\|$	0.2197	0.2023	1.9570	8.9483	442.00
	(0.0038)	(0.0054)	(0.1292)	(0.9769)	[0.00]
R_t^2	0.0892	0.1884	5.9460	56.2030	397.00
	(0.0035)	(0.0132)	(0.5524)	(9.5612)	[0.00]

Japanese Yen

	Mean	Standard Deviation	Skewness	Excess Kurtosis	Q(12)
R_t	-0.0082	0.7313	-0.5871	7.4230	20.30
	(0.0138)	(0.0175)	(0.1971)	(1.0966)	[0.06]
$\|R_t\|$	0.5187	0.5156	2.4366	13.2530	457.00
	(0.0097)	(0.0170)	(0.2534)	(2.8870)	[0.00]
R_t^2	0.5349	1.3583	9.2349	151.7080	396.00
	(0.0257)	(0.1578)	(2.0228)	(55.3563)	[0.00]

Summary statistics for daily exchange rate returns from January 2, 1990 to Dec 29, 2000. Q(12) are modified Ljung-Box Statistics robust to heteroskedasticity for serial correlation with 12 lags. Standard errors robust to heteroskedasticity are in parenthesis, and p-values are in square brackets.

Table 3.2: Estimates of the Constant Jump Models

$$R_t = \mu + \phi R_{t-1} + \sqrt{h_t} z_t + \sum_{k=1}^{n_t} Y_{t,k}$$

$$Y_{t,k} \sim NID(\theta, \delta^2), \quad h_t = \omega + \alpha \epsilon_{t-1}^2 + \beta h_{t-1}$$

$$P(n_t = j) = \frac{\exp^{-\lambda} \lambda^j}{j!}, \; z_t \sim NID(0,1)$$

Parameter	Canada	Japan
μ	0.0059	-0.0106
	(0.0050)	(0.0126)
ϕ	0.0587	-0.0146
	(0.0189)	(0.0182)
ω	0.0002	0.0006
	(0.0006)	(0.0010)
α	0.0581	0.0253
	(0.0082)	(0.0048)
β	0.9171	0.9564
	(0.0089)	(0.0086)
δ	0.3311	0.9157
	(0.0302)	(0.1034)
θ	0.0515	-0.2669
	(0.0252)	(0.0877)
λ	0.2142	0.2056
	(0.0516)	(0.0610)
Q	19.33	19.36
	[0.08]	[0.08]
Q^2	8.80	14.33
	[0.71]	[0.28]
lnL	-347.82	-2785.65

Standard errors are in parentheses. p-values are in square brackets. Q^2 is the modified Ljung-Box portmanteau test, robust to heteroskedasticity, for serial correlation in the squared standardized residuals with 12 lags for the respective models.

Table 3.3: Bivariate BEKK-GARCH Model

BEKK Parameters	Canada		Japan		Canada & Japan
μ_1	0.0043 (0.0049)	μ_2	-0.0032 (0.0123)	c_{12}	0.0192 (0.0085)
ϕ_1	0.0666 (0.0195)	ϕ_2	0.0303 (0.0196)	a_{12}	-0.0125 (0.0072)
c_1	-0.0297 (0.0055)	c_2	-0.0909 (0.0134)	a_{21}	-0.0210 (0.0480)
a_1	0.2277 (0.0175)	a_2	0.1946 (0.0169)	b_{12}	0.0051 (0.0024)
b_1	0.9667 (0.0055)	b_2	0.9730 (0.0050)	b_{21}	0.0077 (0.0135)
Q_1^2	9.45 [0.66]	Q_2^2	13.03 [0.36]	Q_{12}	17.79 [0.12]
lnL	-3324.032				

Standard errors are in parentheses. p-values are in square brackets. Q^2 and Q_{12} are the modified Ljung-Box portmanteau tests, robust to heteroskedasticity, for serial correlation in the squared standardized residuals and the cross products of the standardized residuals with 12 lags.

Table 3.4: Estimates of Correlated Bivariate Poisson GARCH (CBP-GARCH) Model

BEKK Parameters	Canada		Japan		Canada & Japan
μ_1	0.0053 (0.0050)	μ_2	-0.0102 (0.0127)	c_{12}	0.0025 (0.0109)
ϕ_1	0.0581 (0.0190)	ϕ_2	-0.0072 (0.0179)	a_{12}	-0.0076 (0.0067)
c_1	-0.0159 (0.0068)	c_2	-0.0325 (0.0154)	a_{21}	-0.0315 (0.0403)
a_1	0.2364 (0.0187)	a_2	0.1374 (0.0139)	b_{12}	0.0028 (0.0022)
b_1	0.9577 (0.0065)	b_2	0.9821 (0.0037)	b_{21}	0.0101 (0.0128)
Jump Parameters					
λ_1	0.1415 (0.0589)	λ_2	0.1141 (0.0517)	λ_3	0.0861 (0.0329)
θ_1	0.0526 (0.0243)	θ_2	-0.2784 (0.0897)	ρ_{12}	-0.2020 (0.1344)
δ_1	0.3381 (0.0370)	δ_2	0.9477 (0.1038)	$\mathrm{Corr}(n_{1t}, n_{2t})$	0.4033
Q_1^2	18.43 [0.10]	Q_2^2	15.46 [0.21]	Q_{12}	19.76 [0.07]
Q_{ξ_1}	13.84 [0.31]	Q_{ξ_2}	16.56 [0.16]	Q_{ξ_3}	25.45 [0.01]
lnL	-3119.64	LR(7)	408.74 [0.00]		

Standard errors are in parentheses. p-values are in square brackets. Q^2 and Q_{12} are the modified Ljung-Box portmanteau tests, robust to heteroskedasticity, for serial correlation in the squared standardized residuals and the cross products of the standardized residuals with 12 lags. Q_{ξ_i} is the same test for serial correlation in the jump intensity residuals. LR(7) is a non-standard test of jumps which is chi square distributed with df=7.

Table 3.5: Estimates of Correlated Bivariate Poisson GARCH Model (CAN/JY) with Time Varying Intensities (CBP-GARCH-R^2)

BEKK Parameters	Canada		Japan		Canada & Japan
μ_1	0.0057 (0.0050)	μ_2	-0.0084 (0.0126)	c_{12}	0.0035 (0.0090)
ϕ_1	0.0567 (0.0194)	ϕ_2	-0.0114 (0.0189)	a_{12}	-0.0035 (0.0065)
c_1	-0.0140 (0.0070)	c_2	-0.0354 (0.0141)	a_{21}	-0.0348 (0.0415)
a_1	0.2160 (0.0214)	a_2	0.1242 (0.0138)	b_{12}	0.0023 (0.0021)
b_1	0.9625 (0.0069)	b_2	0.9834 (0.0036)	b_{21}	0.0110 (0.0129)
Time-Varying Jump Parameters					
λ_1	0.1482 (0.0646)	λ_2	0.1157 (0.0555)	λ_3	0.0756 (0.0361)
η_1	0.5712 (0.2979)	η_2	0.1167 (0.2089)	η_3	0.2818 (0.0727)
				η_4	0.3675 (0.2146)
θ_1	0.0452 (0.0209)	θ_2	-0.2345 (0.0749)	ρ_{12}	-0.1757 (0.2395)
δ_1	0.3061 (0.0339)	δ_2	0.8945 (0.0979)		
Q_1^2	14.62 [0.26]	Q_2^2	16.51 [0.16]	Q_{12}	18.68 [0.09]
lnL	-3106.71	LR(4)	25.86 [0.33E-04]		

Standard errors are in parentheses. p-values are in square brackets. Q^2 and Q_{12} are the modified Ljung-Box portmanteau tests, robust to heteroskedasticity, for serial correlation in the squared standardized residuals and the cross products of the standardized residuals with 12 lags. LR(4) is the χ^2 test with four degrees of freedom for the null of constant intensities.

Table 3.6: Estimates of Correlated Bivariate Poisson GARCH Model with Autoregressive Jump Intensities (CBP-GARCH-ARJI)

BEKK Parameters	Canada		Japan		Canada & Japan
μ_1	0.0060 (0.0049)	μ_2	-0.0077 (0.0123)	c_{12}	-0.0049 (0.0108)
ϕ_1	0.0574 (0.0190)	ϕ_2	-0.0111 (0.0183)	a_{12}	0.0018 (0.0056)
c_1	-0.0067 (0.0107)	c_2	-0.0188 (0.0190)	a_{21}	0.0446 (0.0498)
a_1	0.1634 (0.0249)	a_2	0.1044 (0.0134)	b_{12}	0.0002 (0.0014)
b_1	0.9775 (0.0065)	b_2	0.9879 (0.0029)	b_{21}	-0.0118 (0.0119)
ARJI Jump Parameters					
λ_1	0.1999 (0.0658)	λ_2	0.2052 (0.0732)	λ_3	0.0177 (0.0020)
ζ_1	6.0E-09 (0.0275)	ζ_2	1.89E-10 (0.0126)	ζ_3	0.9103 (0.0335)
γ_1	0.0388 (0.1748)	γ_2	0.3441 (0.2746)	γ_3	0.5278 (0.1582)
θ_1	0.0357 (0.0176)	θ_2	-0.1890 (0.0506)	ρ_{12}	-0.1551 (0.0862)
δ_1	0.2883 (0.0300)	δ_2	0.7542 (0.0869)		
Q_1^2	12.55 [0.40]	Q_2^2	18.71 [0.09]	Q_{12}	16.85 [0.15]
Q_{ξ_1}	11.40 [0.49]	Q_{ξ_2}	13.01 [0.36]	Q_{ξ_3}	11.84 [0.45]
lnL	-3098.22	LR(6)	42.84 [5.71E-08]		

Standard errors are in parentheses. p-values are in square brackets. Q^2 and Q_{12} are the modified Ljung-Box portmanteau tests, robust to heteroskedasticity, for serial correlation in the squared standardized residuals and the cross products of the standardized residuals with 12 lags. Q_{ξ_i} is the same test for serial correlation in the jump intensity residuals. LR(6) is the χ^2 test statistics with six degrees of freedom for the null of constant jump intensities.

Table 3.7: RMSE and MAE of Conditional Volatility Forecasts

RMSE					
	BM	BEKK	CBP	CBP-R^2_{t-1}	CBP-ARJI
In-Sample					
Canada	0.2087	0.2901	0.1970	0.1960	0.1956
Japan	0.5312	0.5306	0.4961	0.4956	0.4964
Out-of-Sample					
Canada	0.3457	0.2578	0.2479	0.2461	0.2440
Japan	0.5810	0.4212	0.3909	0.3911	0.3942
MAE					
	BM	BEKK	CBP	CBP-R^2_{t-1}	CBP-ARJI
In-Sample					
Canada	0.4337	0.1674	0.1477	0.1456	0.1438
Japan	0.6825	0.4202	0.3574	0.3532	0.3468
Out-of-Sample					
Canada	0.5103	0.2177	0.2029	0.1994	0.1943
Japan	0.6776	0.3615	0.3167	0.3146	0.3098

RMSE and MAE refer to root mean square errors and mean absolute errors, respectively. BM represents the benchmark forecast using last period's true standard deviation. CBP, CBP-R^2_{t-1}, and CBP-ARJI correspond to the correlated bivariate Poisson GARCH models with constant jump intensity, time-varying jump intensity conditioning on last period's volatility, and autoregressive conditional jump intensity.

Table 3.8: RMSE and MAE of Conditional Volatility Forecasts with R_{t-1}^2 as True Variance

RMSE					
	BM	BEKK	CBP	CBP-R_{t-1}^2	CBP-ARJI
In-Sample					
Canada	0.2641	0.2087	0.1970	0.1960	0.1876
Japan	0.6654	0.5316	0.4964	0.4960	0.4862
Out-of-Sample					
Canada	0.3465	0.2582	0.2489	0.2470	0.2353
Japan	0.5807	0.4214	0.3907	0.3909	0.3872
MAE					
	BM	BEKK	CBP	CBP-R_{t-1}^2	CBP-ARJI
In-Sample					
Canada	0.4338	0.4087	0.3843	0.3815	0.3727
Japan	0.6824	0.6488	0.5979	0.5944	0.5844
Out-of-Sample					
Canada	0.5114	0.4669	0.4505	0.4464	0.4328
Japan	0.6777	0.6016	0.5630	0.5613	0.5525

The true standard deviation is approximated by $\sqrt{R_{t-1}^2}$. RMSE and MAE refer to root mean square errors and mean absolute errors, respectively. BM represents the benchmark forecast using last period's true standard deviation. CBP, CBP-R_{t-1}^2, and CBP-ARJI correspond to the correlated bivariate Poisson GARCH models with constant jump intensity, time-varying jump intensity conditioning on last period's volatility, and autoregressive conditional jump intensity.

Figure 3.1: Canadian Dollar Spot Rates (A) and Returns (B) 2/1/90-29/12/00

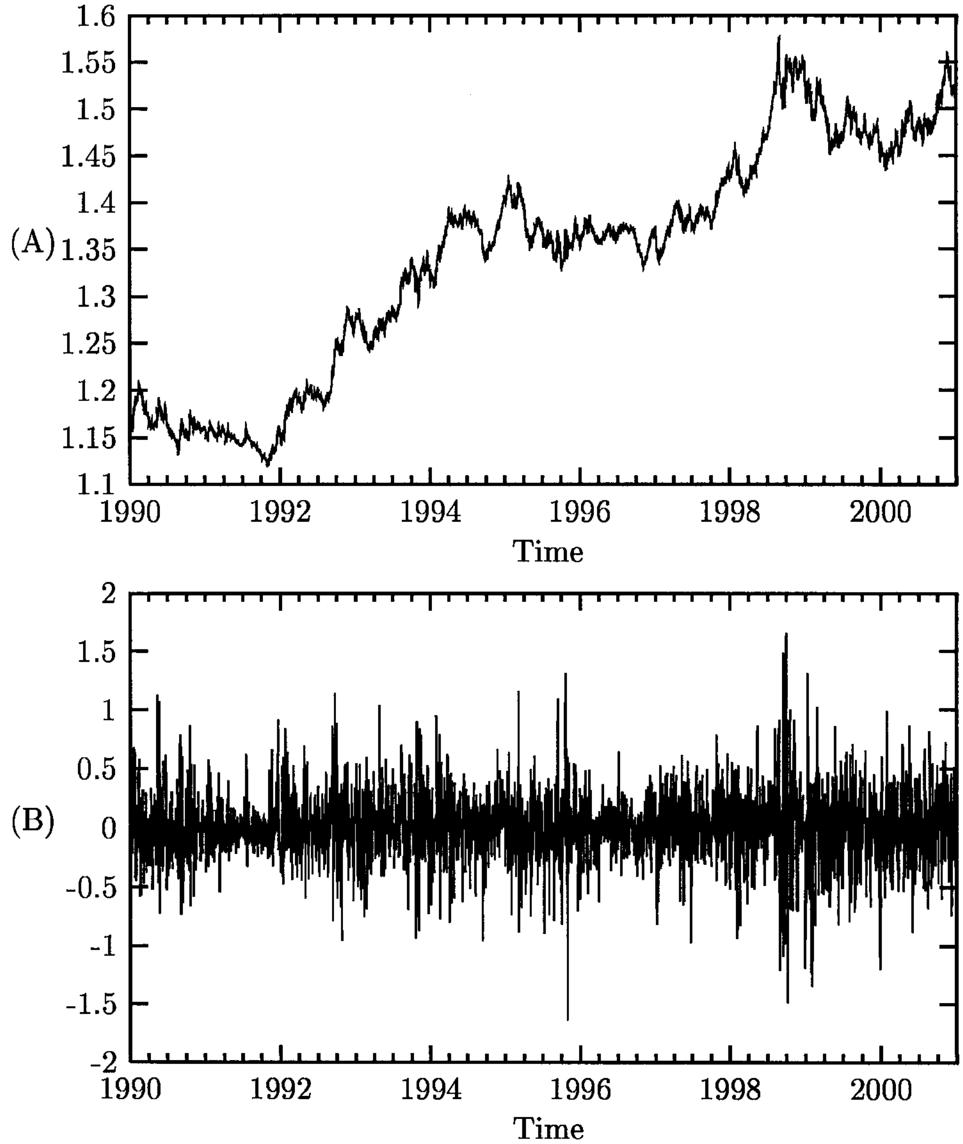


Figure 3.2: Japanese Yen Spot Rates (A) and Returns (B) 2/1/90-29/12/00

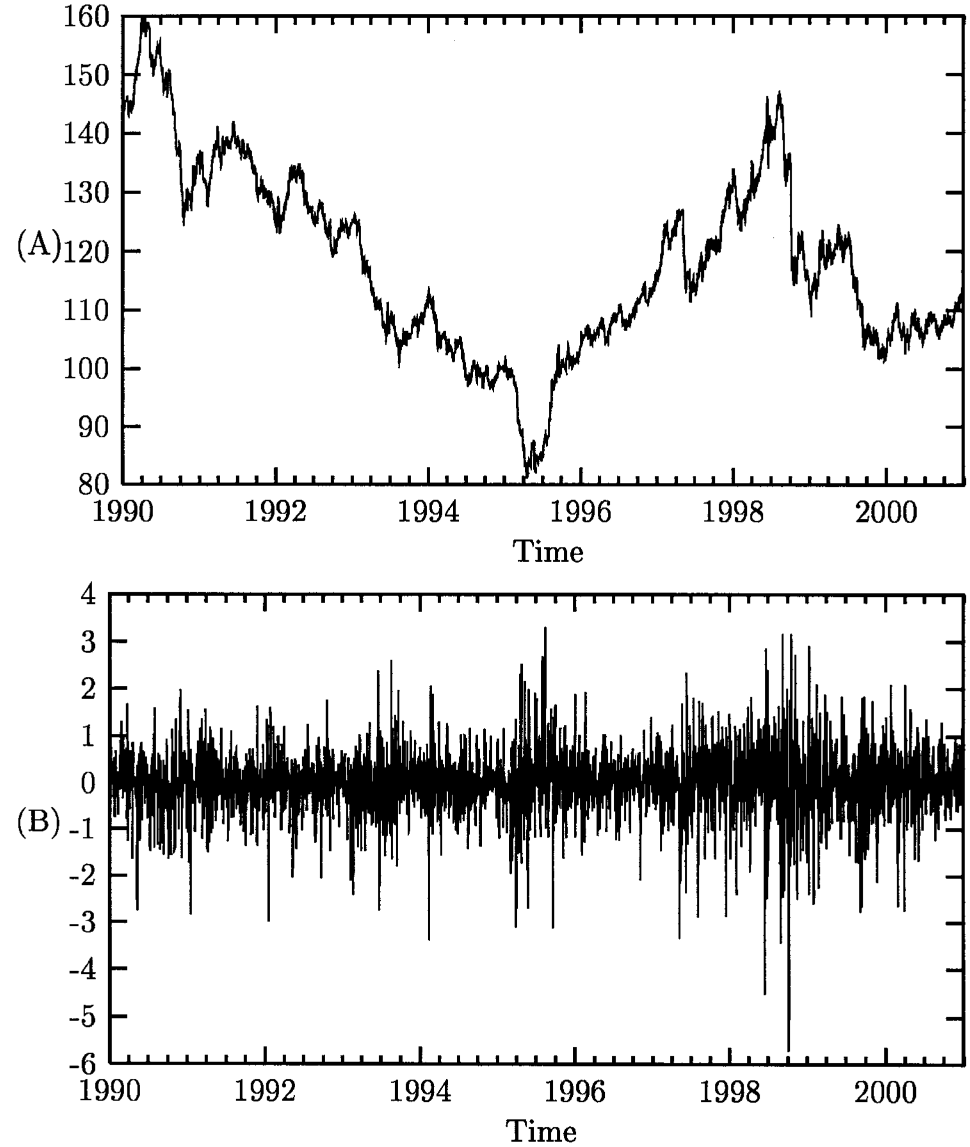


Figure 3.3: Time Varying Jump Intensities from CBP-GARCH-R^2_{t-1} Model

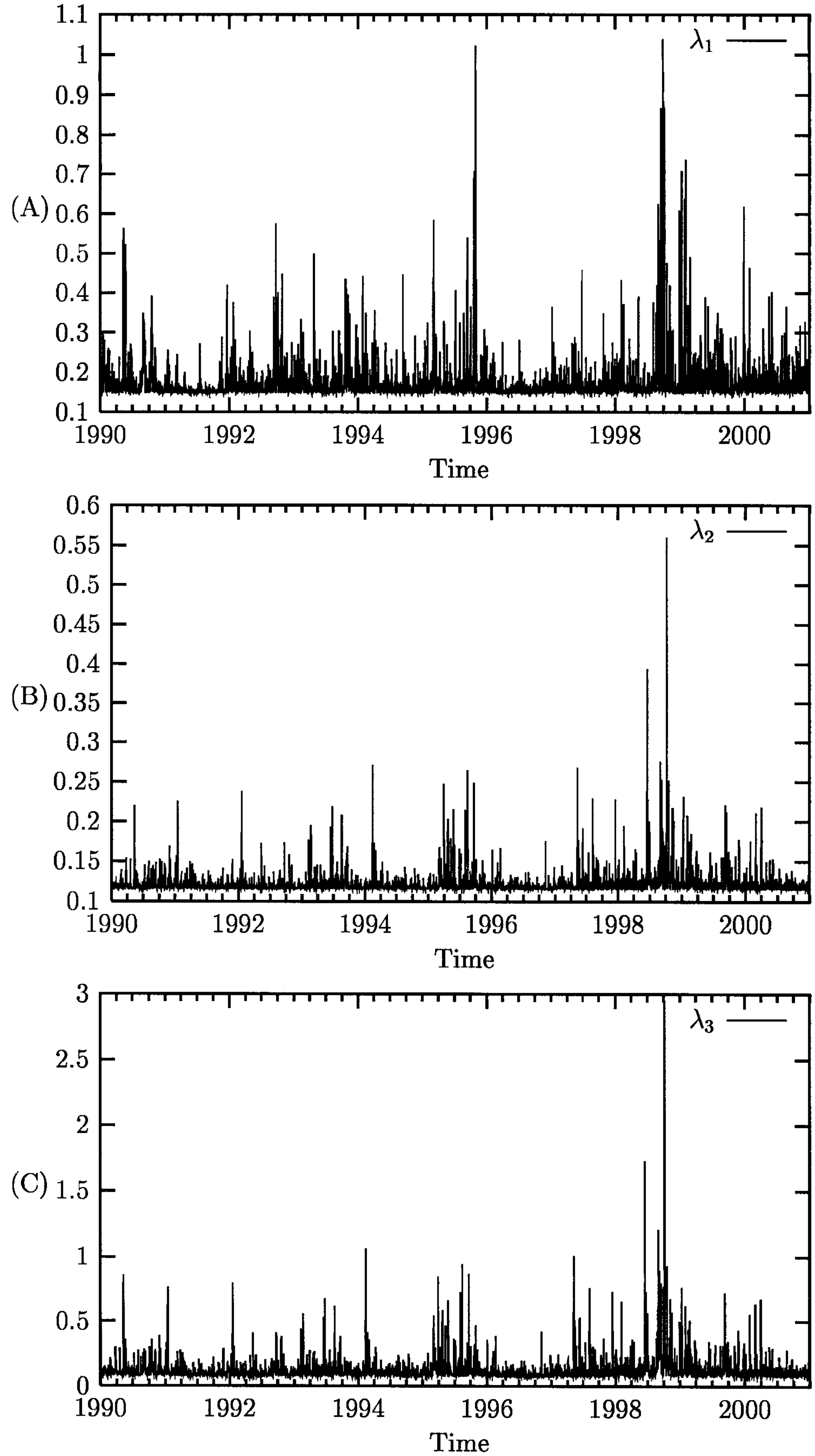


Figure 3.4: Time Varying Jump Intensities from CBP-GARCH-ARJI Model

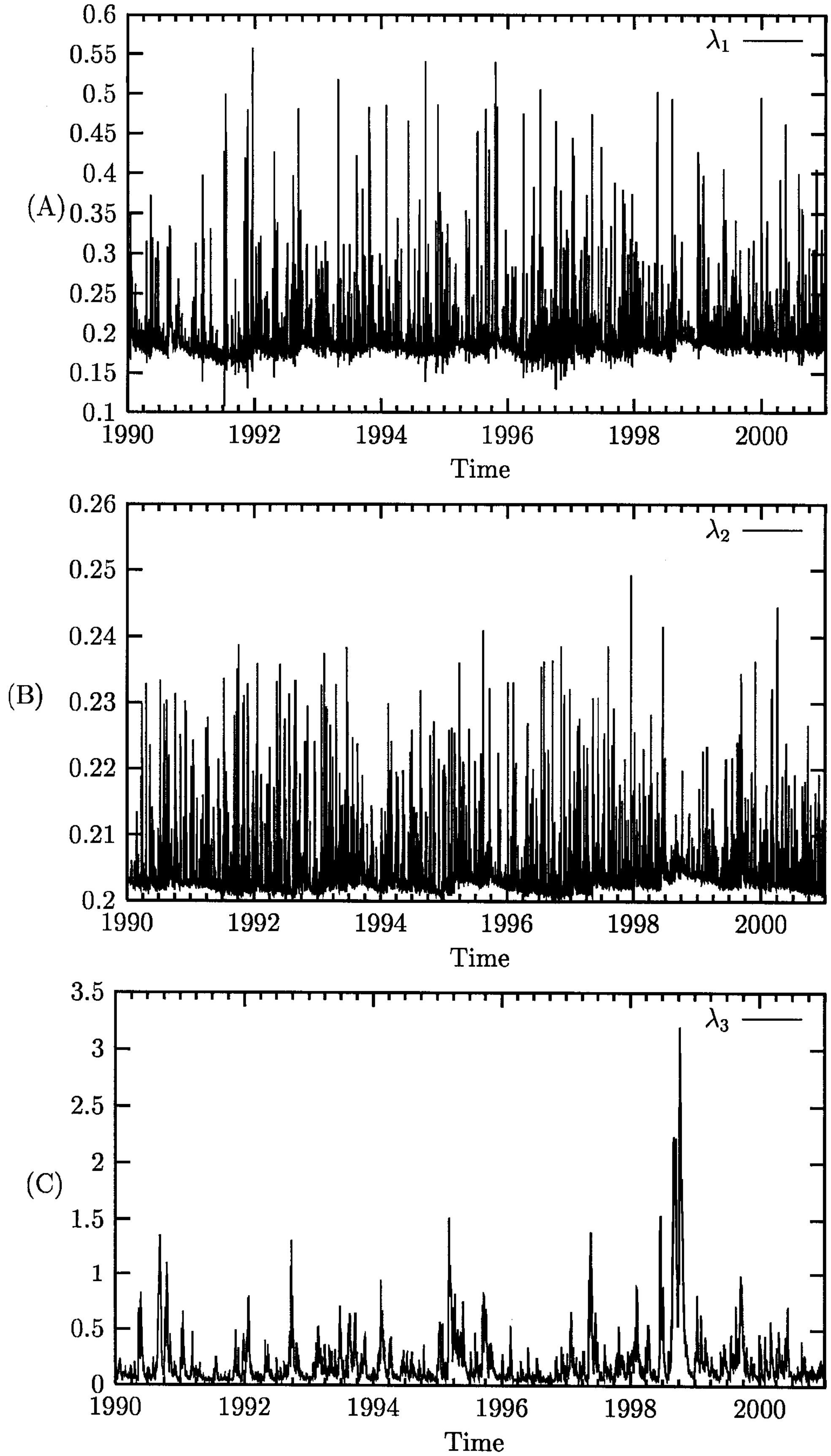
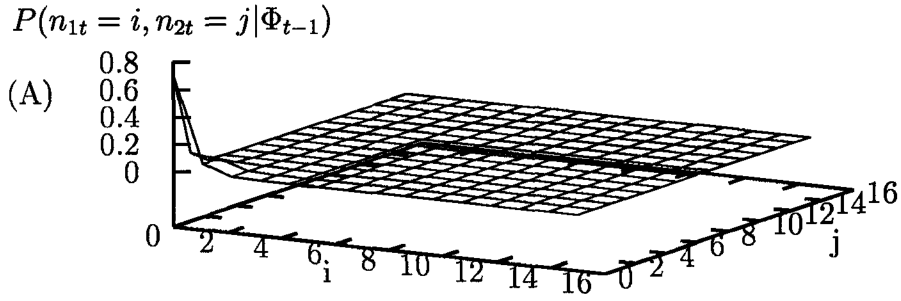
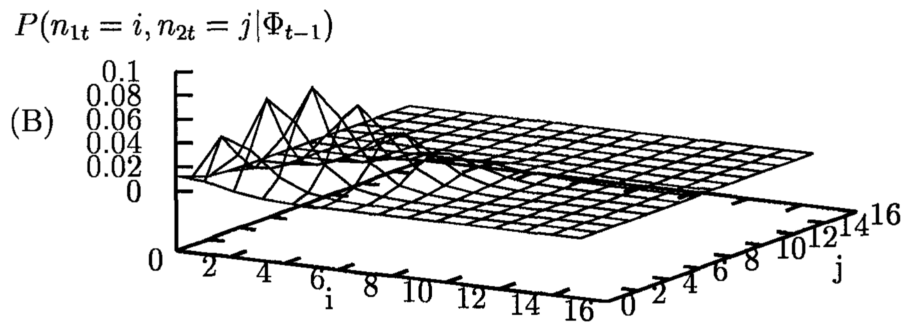


Figure 3.5: Conditional Probabilities of Jumps

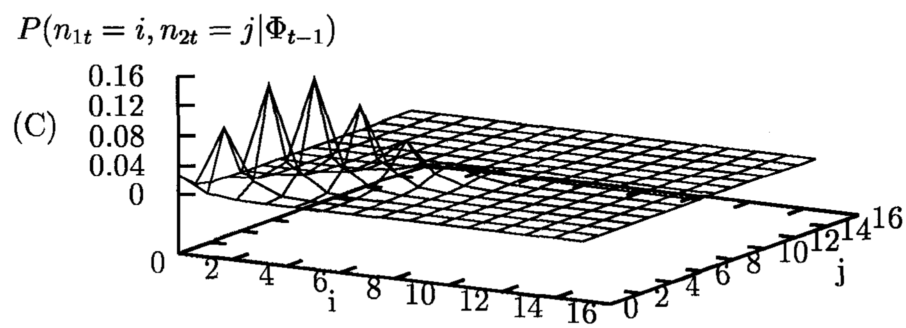
Constant λ_i ($\lambda_1 = 0.1141, \lambda_2 = 0.1415, \lambda_3 = 0.0861$)



October 8, 1998 ($\lambda_1 = 0.5699, \lambda_2 = 0.8696, \lambda_3 = 2.9641$)

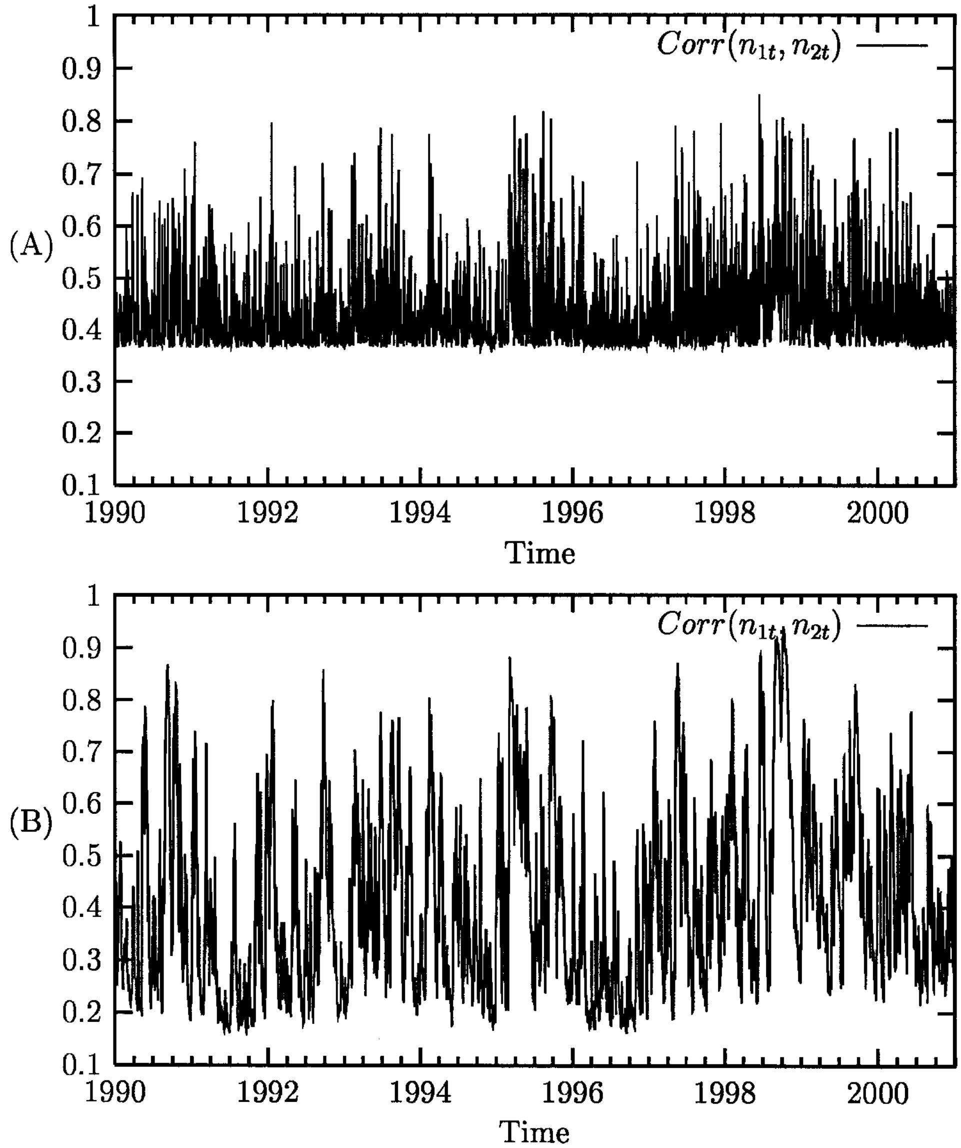


October 8, 1998 ($\lambda_1 = 0.2098, \lambda_2 = 0.2105, \lambda_3 = 3.1985$)



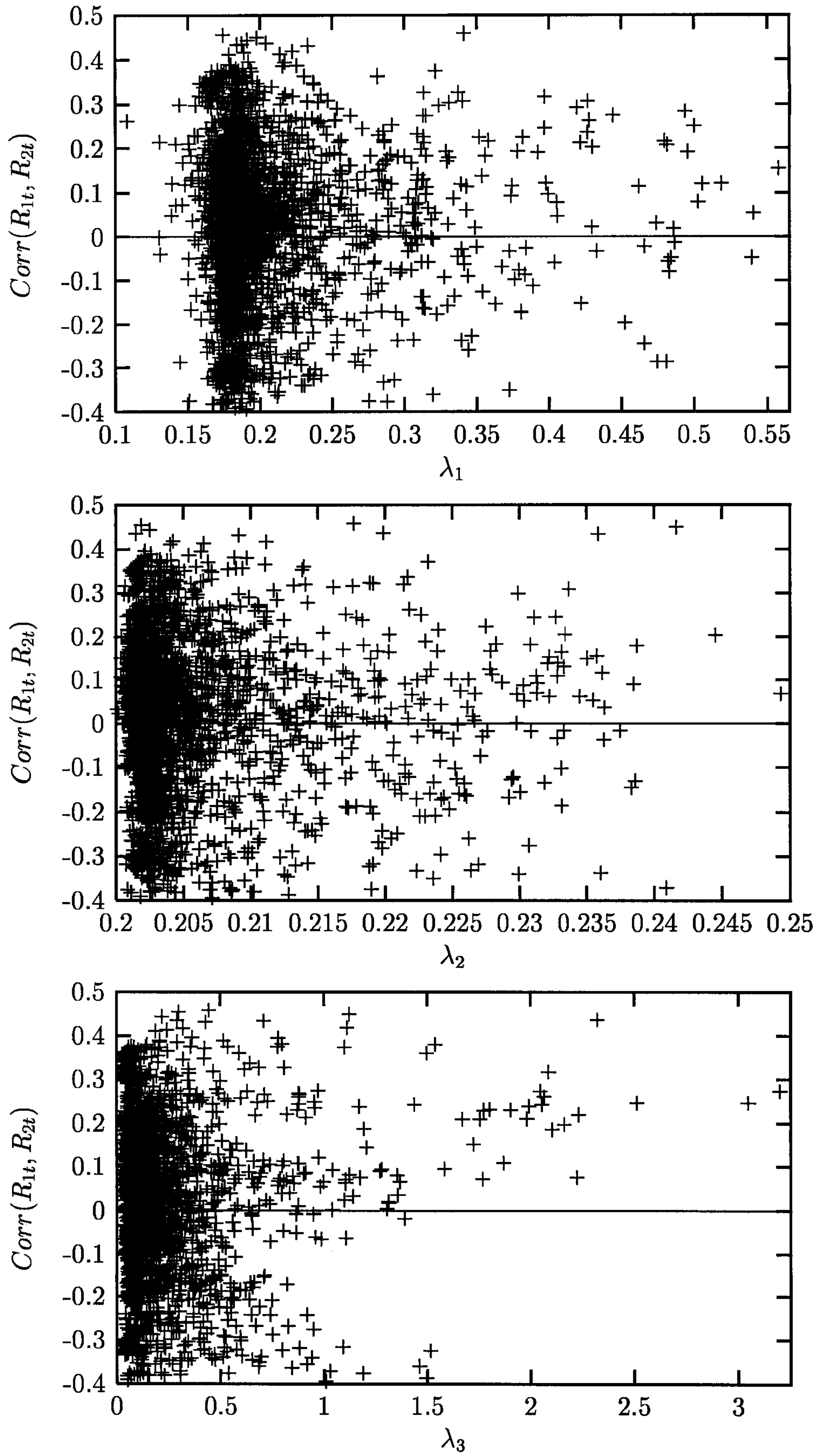
Note: (A) The CBP-GARCH Model with constant jump intensities. (B) The CBP-GARCH-R^2_{t-1} Model. (C) The CBP-GARCH-ARJI Model.

Figure 3.6: Correlations between Jump Frequencies



Note: (A) The CBP-GARCH-R^2_{t-1} Model. (B) The CBP-GARCH-ARJI Model.

Figure 3.7: Time Varying Jump Intensities and Correlations between Returns



Chapter 4

Time-Varying Risk Premia in Nonrenewable Resource Asset Returns

4.1 Introduction

Recently, the effect of introducing uncertainty into the Hotelling model has been considered in a number of empirical studies. For example, Slade and Thille (1995) examined the risk-adjusted continuous-time Hotelling model based on the theoretical work of Gaudet and Khadr (1991) and found that the Hotelling rule can not be rejected. Young and Ryan (1996) estimated the Euler equations arising from Gaudet and Howitt (1989) that yield the risk-adjusted "Hotelling" rule, and based on Hansen's J-test, found that the data were not inconsistent with the model. Using Generalized Method of Moments (GMM) estimation they showed that a significant risk premium exists at the industrial-level data for certain metals (copper, lead, silver, and zinc). The risk premia that they calculated were treated as constants across the 40 year span of the annual data.

The assumption of a constant risk premium is most likely not appropriate for metal prices. Slade (1982) demonstrates that metal prices have experienced high fluctuations over the period 1870-1979 which is reflected by many of the metal prices following a U-Shaped pattern. Since agents perception of uncertainty (risk) may depend on the price movement, the risk premium demanded by the agents will change over time depending on the recent volatility patterns of metal prices.

Estimation of a time-varying risk premium model not only provides potential insight into the failure of the Hotelling rule in empirical studies, but may also improve the predictive power of existing models. For example, Heal and Barrow (1980) studied the relationship between metal prices and interest rates with the no arbitrage model ignoring the risk premium. Explicitly allowing for risk premia in the context of a general dynamic model may yield interesting general insights into the behaviour of the metal prices.

In attempts to fully capture all of the dynamics, several studies have found the presence of systematic jumps in natural resource prices (Khalaf, Saphores, and Bilodeau (2000), Saphores, Khalaf, and Pelletier (2000)). The occurrence of jumps is motivated by changes of unusual magnitude that follow the arrival of news unanticipated by market participants. As the market absorbs this information, a series of abrupt changes may be observed in the return series. This characteristic contributes to the return series having a fat tailed distribution. Ignoring such excess kurtosis may lead to mispricing of derivatives which rely on the underlining financial asset.

The purpose of this research is to consider the presence of time-varying risk premia in resource asset returns. To properly measure the time varying risk premia, we use a bivariate GARCH-M model based on the consumption-based Intertemporal Asset Pricing Model (IAPM). In this case, the risk premia are modeled in relation to a benchmark portfolio assumed to be perfectly correlated with marginal utilities in different time periods. As an extension to the model, we also consider the possibility of systematic jumps simultaneously affecting the metal prices and benchmark portfolio.

These models allow us to formally test hypotheses concerning the influence of time varying risk premia on returns. The risk premia are modeled as proportional to the conditional covariance between metal prices and the benchmark portfolio. Allowing the ratio of covariance and variance to enter the mean equation in the bivariate model allows us to test for the presence of risk premia. Our examination of time-varying risk premia is performed in the context of the relationship between interest rates and four metal prices (copper, lead, silver, and zinc) currently traded on the London Metal Exchange (LME).

The organization of the chapter is as follows: Section 4.2 reviews some of the literature on the development of the risk-adjusted Hotelling rule. Section 4.3 specifies the models. Section 4.4 discusses estimation procedures and econometric issues related to the GARCH-M models. A description of the data set is provided in Section 4.5. Section 4.6 presents the estimation result, including a discussion of the characteristics of the time-varying risk premia found in the two models, and finally, Section 4.7 concludes.

4.2 Literature Review

The most basic formulation of the Hotelling (1931) rule states that under a set of restrictive conditions (competitive market structure, no uncertainty, and no stock effect) nonrenewable asset returns should be equal to the interest rate. Hotelling formalizes this relationship in terms of the following equation:

$$\frac{\dot{p}_t}{p_t} = r_t \tag{4.1}$$

where $\dot{p}_t$ is the change in resource price, net of marginal cost, at time t and r_t is the interest rate at time t. This simple Hotelling rule has motivated a voluminous literature that studies the relationship between interest rates and metal prices.

Many empirical studies of the relationship between resource prices and interest rates have been undertaken. Heal and Barrow (1980) test a series of models based on the idea that the return on a nonrenewable resource affects the holding decisions between resource markets and markets of other assets. Their arbitrage models examine the relationship between the interest rate and metal prices and conclude that it is the change in the interest rate, rather than the level, that affects price changes.

Agbeyegbe (1989) employs Hendry's general to specific strategy in the context of a simple expectations model to investigate the relationship between the interest rate and metal prices. A dynamic model, involving lagged dependent variables and first differenced interest rates, shows that the differenced interest rates are important determinants of metal price movements. These results lend support to the findings of Heal and Barrow. Smith (1981) compares five different models including the simple Hotelling model, a multiple rate model, a general lag model, an Almon lag model, and the Heal and Barrow model. Based on the forecasting performance (RMSE, MAE and Theil's Inequality Coefficients), Smith concludes that Heal and Barrow's model with the differenced interest rates instead of the level of the interest rate best fits the twelve examined minerals over the period of 1900 through 1973. These studies conclude that the arbitrage type of models like those of Heal and Barrow (1980) and Agbeyegbe (1989) are superior to the simple Hotelling model in terms of explaining movements.

Subsequent to these results, another stream of the theoretical resource literature emerged, explicitly introducing uncertainty into Hotelling-type models of optimal resource extraction.[1] Gaudet and Khadr (1991) derive a stochastic version of the Hotelling rule based on an equilibrium condition in an intertemporal asset-pricing model. Uncertainty is introduced into the model through stochastic indices in the production and extraction processes. The stochastic Hotelling rule is very similar to the results of the standard Capital Asset Pricing Model (CAPM) in which the excess

[1]Other recent extensions to the Hotelling model are discussed in Krautkraemer (1998)

return of the resource is equal to the product of the beta coefficient and the excess return of the chosen portfolio. Slade and Thille (1995) estimate the model of Gaudet and Khadr using panel data from a set of Canadian copper mines. These firm-level data provide important information on extraction cost. Their results indicate that the hypothesis of a stochastic Hotelling rule based on the CAPM can not be rejected.

Similar to Gaudet and Khadr (1991), Gaudet and Howitt (1989) modify the simple Hotelling rule to account for uncertainty on the productivity of capital and of the natural resource in the context of a simple 2-period macro model. The equilibrium conditions indicate that the expected rate of increase of the net resource price should equal the expected interest rate, plus a risk premium which is a function of the covariance between the marginal utilities of consumption and the rate of net price increase. The sign and size of this covariance term (as the risk premium) is an empirical question. Young and Ryan (1996) estimate the Gaudet and Howitt model using Generalized Method of Moments (GMM) and find that the incorporation of a constant risk premium into the Hotelling equations improves the performance of their industry-level Hotelling model. Sensitivity analysis was performed with two utility specifications (Constant Relative Risk Aversion (CRRA) and Constant Absolute Risk Aversion (CARA)), two interest rates (30-day Treasury bills and Government Bonds), and net prices versus gross prices. They did not, however, allow for any variation over time in the risk premium. The risk premium calculated in this industry-level study was constructed based on estimates of marginal utility using annual consumption data.

Alternatively, we can impose further structure on risk premia to avoid the requirement of collecting data on consumption. We follow the methodology used in the foreign exchange literature (Hansen and Richard (1987), Hansen and Hodrick (1983), McCurdy and Morgan (1991,1992, and 1993)) where it is assumed that there is a benchmark portfolio correlated with the risk premia so that a benchmark index can be substituted for the consumption data. In addition, this bivariate GARCH-M model allows the risk premia to vary over time which relaxes the assumption of constant risk premium in Young and Ryan (1996).

This approach can be augmented with a systematic jump component to model the leptokurtosis often observed in financial data series. Systematic jumps in resource prices is not a new idea, and these abrupt changes are documented in many empirical studies. For example, Saphores, Khalaf, and Pelletier (2000) report the presence of systematic jumps in time series of stumpage prices based on a Monte Carlo Simulation test. Khalaf, Saphores, and Bilodeau (2000) investigate and confirm jump dynamics in weekly spot prices of copper and nickel.

The approach, outlined in further detail in Section 4.4, allows for the use of higher frequency data. Thus, as an extension to the current literature on uncertainty, we incorporate a time-varying risk premium into the basic Hotelling framework in the context of bivariate GARCH-M model. The risk premium is modeled by the use of a benchmark portfolio. Before describing the econometric methodology and the data, we present a basic outline of the theoretical model of Gaudet and Howitt (1989).

4.3 Model

This section sets out the theoretical results underlying our empirical work. We will first discuss the risk-adjusted Hotelling rule (with risk premium) proposed by Gaudet and Howitt (1989) in a 2-period consumption model. Then we modify this rule in terms of the marginal rate of substitution (i.e., the discounted ratio of marginal rates of utilities) to motivate the Intertemporal Asset Pricing Model (IAPM).

4.3.1 Risk-Adjusted Hotelling Rules

Gaudet and Howitt (1989) consider a 2-period competitive economy which produces a single capital consumption good. They introduce uncertainty, as a technological shock in the second period, to the standard Hotelling framework which, under the assumption of competitive markets, produces the following first order conditions:

$$-U'(C_1) + \beta EU'(C_2) \cdot [1+\rho] = 0 \tag{4.2}$$

$$-U'(C_1) + \beta EU'(C_2) \cdot [1+r] = 0 \tag{4.3}$$

$$U'(C_1) \cdot (P_1) - \beta EU'(C_2) \cdot (P_2) = 0 \tag{4.4}$$

P_i (i=1,2) is the relative net resource price at period i; ρ is the rate on an alternative risk-free bond; $U'(C_i)$ is the marginal utility of consumption in period i; r is the rate of return on an alternative risky capital asset; β is the discount factor; and E is the expectation operator.

The risk-adjusted Hotelling rules derived from equations (4.2) to (4.4) are

$$E\left(\frac{\Delta P}{P}\right) = \rho - \frac{Cov(U'(C_2), \Delta P/P)}{E(U'(C_2))} \tag{4.5}$$

$$E\left(\frac{\Delta P}{P}\right) = E(r) - \frac{Cov(U'(C_2), (\Delta P/P) - r)}{E(U'(C_2))}. \tag{4.6}$$

Equation (4.5) describes the equilibrium relationship between resource prices and the return on a risk-free bond, with the difference between the expected returns being the risk premium. The sign of the risk premium depends on whether the returns on the resource tend to be high (or low) during periods of high marginal utility (low consumption). In the latter case, for example, the resource asset must offer the agent an additional risk premium of $-\frac{Cov(U'(C_2),\Delta P/P)}{E(U'(C_2))}$. From equation (4.6) we see that when both assets are risky, it is their relative riskiness that determines the sign and size of the risk premium.

4.3.2 The Intertemporal Asset Pricing Model (IAPM)

The risk-adjusted Hotelling rules can also be rewritten in terms of the marginal rate of substitution[2] M_s as

$$E\left(\frac{\Delta P}{P}\right) = \rho - \frac{Cov(M_s, \Delta P/P)}{E(M_s)} \tag{4.7}$$

$$E\left(\frac{\Delta P}{P}\right) = E(r) - \frac{Cov(M_s, \Delta P/P - r)}{E(M_s)} \tag{4.8}$$

Estimation of equations (4.7) and (4.8) or of the first order conditions that yield these rules requires data on consumption which can be difficult to obtain, especially for use with even relatively high frequency resource price data, such as monthly data sets. One way to overcome this problem, is to adopt the approach from the foreign exchange literature, (Hansen and Richard (1987), Hansen and Hodrick (1983), and McCurdy and Morgan (1991, 1992, and 1993)), in which the first order conditions are defined as a consumption-based Intertemporal Asset Pricing Model (IAPM). In this application of the IAPM model, the nonrenewable resource is treated as one of the financial assets.

To circumvent the problem of constructing a marginal utility variable (as a function of consumption), we assume that there is a benchmark portfolio perfectly correlated with M_s.[3] As a result, all of the M_s terms in equations (4.7) and (4.8) can be replaced by the return of that portfolio. Intuitively, preferences related to having more consumption in the first period than the second period depends on the opportunities given by the benchmark portfolio. It may be optimal to give up some consumption now because a high yield from the benchmark portfolio is expected next period. To introduce this portfolio formally, a benchmark portfolio with the return

[2] See Appendix for the derivation of equations (4.7) and (4.8).

[3] A weaker condition is developed by Breeden, Gibbons, and Litzenberger (1989) in which the benchmark portfolio is maximally correlated with the marginal rate of substitution instead of perfectly conditionally correlated. The maximal correlated portfolio is defined as the portfolio with the strongest correlation with the marginal rate of substitution compared to any other portfolio.

R_{Bt} is defined as a linear combination of R_{Ct} and a risk-free bond, where R_{Ct} is the return of the portfolio which is perfectly conditionally correlated with the marginal rate of substitution. The benchmark portfolio will be conditionally mean-variance efficient.[4]

Given these characteristics, the expected return on any asset will follow the standard conditional CAPM equilibrium in which the return will be a function of the conditional covariance, variance, and benchmark portfolio. Equation (4.7) can be re-expressed in terms of this conditional beta asset pricing relation, where the covariance between the marginal rate of substitution and risk-free rate is replaced by the covariance between benchmark returns and the risk-free rate. This relationship can be represented as

$$E_{t-1}\left(\frac{\Delta P_t}{P_t}\right) - \rho = \frac{Cov_{t-1}(R_{Bt}, \Delta P/P)}{Var_{t-1}(R_{Bt})}(E_{t-1}(R_{Bt}) - \rho). \tag{4.9}$$

where the time subscripts 't-1' have been added to the covariance and variance terms to relax the assumption of a constant risk premium.

4.4 Econometric Methodology

The family of the ARCH-M models has been used extensively in the field of financial economics. Studies such as Adams and Moghaddam (1991), Fischer (1988), Lee (1988), and Lee and Tse (1991) successfully apply the ARCH-M model to stock prices, bond prices, and exchange rates to capture the variation associated with the conditional variance. A detailed review of the use of ARCH model in finance can be found in Bollerslev, Chou, and Kroner (1992), Bollerslev, Engle, and Nelson (1994), and Higgins and Bera (1992).

[4]Conditionally mean-variance efficiency refers to the fact that any other portfolio with the same rate of returns will have a higher variance (risk).

4.4.1 Univariate GARCH-M model

The classical univariate GARCH-M model allows for the direct incorporation of a time-varying risk premium into an easily estimated model. It can be written as

$$Y_{it}|X_{it} \sim N(\beta' X_{it} + \delta h_{it}, h_{it}^2), \tag{4.10}$$

$$h_{it}^2 = \alpha_0 + \alpha_1 \epsilon_{it-1}^2 + \gamma_1 h_{it-1}^2 \tag{4.11}$$

$$\epsilon_{it-1} = Y_{it-1} - \beta' X_{it-1} - \delta h_{it-1} \tag{4.12}$$

where, in the context of a risk-adjusted Hotelling rule, Y_{it} (i=1,...,4)[5] represents the nonrenewable asset return, $\beta' X_{it}$[6] captures the effects of the interest rate on the resource return, and δh_{it} captures the risk premium which, since it is a function of the volatility of the resource price, is allowed to vary across time. In the spirit of Heal and Barrow (1980) and Agbeyegbe (1989), a more general dynamic specification of the relationship between resource returns and interest rates can easily be accommodated by expanding X_{it} to include lagged interest rates and resource returns. Note that equation (4.11) contains a standard GARCH parameterization of the variance of the resource return. The restriction $\alpha_1 + \gamma_1 < 1$ must hold to ensure that the unconditional variance of the return is finite.

This model allows us to formally conduct hypothesis tests concerning the presence of a time varying risk premium on returns. In the spirit of Merton (1980)'s model, the risk premium δh_t can be viewed as capturing the effects of the second term in equation (4.5)

$$\frac{Cov_{t-1}(U'(C_2), \Delta P/P)}{E_{t-1}(U'(C_2))}$$

This univariate GARCH-M model avoids the problem of constructing data on $U'(C_2)$ by assuming the risk premium is moving with the conditional standard deviation of

[5] i is used to index the individual resources. (copper, lead, silver, and zinc)

[6] For the simple Hotelling rule posted in equation (4.1), X_{it} would be a scalar (the interest rate), the corresponding β would equal one, and δ would equal zero.

the resource return. Assuming that the risk premium in equation (4.5) is proportional to the volatility of the asset returns, using the conditional heteroskedastic standard deviation, h_t as one of the regressors in the mean equation $(\beta' X_t + \delta h_t)$ can serve as a direct test for the presence of a time varying risk premium. If we cannot reject the null hypothesis that δ=0, this implies that the time-varying risk premium has a significant effect on the resource return. As noted in Gaudet and Howitt (1989), the sign of risk premium is not known *a priori.* It is an empirical matter that depends on the relative riskiness of the resource price and the alternative asset.

However, given that we have no proxy variables to model the change of marginal utilities, solely relying on the conditional variance to capture risk premia may induce mis-specification problem. This motivates the use of the bivariate GARCH model which adapts a benchmark portfolio to model the marginal rate of substitution.

4.4.2 Positive Definite Bivariate GARCH-M Model

A positive definite Bivariate GARCH-M specification is used for the estimation of the IAPM version of the risk-adjusted Hotelling model. Assuming rational expectations, the bivariate statistical model is defined as

$$\begin{aligned} R_{it} &= \gamma_{i0} + \mu_i \frac{h_{iB,t}}{h_{Bt}} (\gamma'_B x_{B,t-1}) + \epsilon_{it}, && (4.13) \\ R_{Bt} &= \gamma'_B x_{B,t-1} + \epsilon_{Bt}, && (4.14) \\ \epsilon_t | I_{t-1} &\sim N(0, H_t). \end{aligned}$$

R_{it} refers to the excess returns on the nonrenewable resource; R_{Bt} represents the excess returns of the benchmark portfolio; x_B is a set of instruments (the lags of R^*_{Bt}) explaining the movement of the benchmark return; h_{iB} represents the conditional covariance between the returns on resource i and the returns of the benchmark portfolio; h_B is the conditional variance of the returns of the benchmark portfolio; and finally γ and μ are parameters.

The time-varying risk premium is captured by the term $\mu h_{iB,t}/h_{B,t}(\gamma_B' x_{B,t-1})$ in the mean equation (4.13). By testing the null hypothesis that μ=0, we can examine the effect of the time-varying risk premium on excess returns. This model can be viewed as a bivariate GARCH-in-Mean model in which both the conditional variance and covariance enter the mean equation.

The choice of the benchmark portfolio is very important as the time-varying risk premia consist of the covariances between the rates of return on the metals and the benchmark portfolio. The risk premia will not be found unless a benchmark correlated with the metals is used. For the extreme case, when the covariance between the benchmark portfolio and the excess returns is zero, the risk premium disappears in equation (4.9), resulting in the simple Hotelling rule.

Similar to the univariate model, certain restrictions are required to ensure stationarity of the model. Therefore, the variance covariance matrix H_t takes the BEKK form as proposed by Baba, Engle, Kraft, and Kroner (1989) to ensure positive definiteness:

$$H_t = CC' + A'\epsilon_{t-1}\epsilon_{t-1}'A + B'H_{t-1}B \tag{4.15}$$

$$\begin{bmatrix} h_{i,t} & h_{iB,t} \\ h_{iB,t} & h_{B,t} \end{bmatrix} = \begin{bmatrix} c_i & 0 \\ c_{iB} & c_B \end{bmatrix}\begin{bmatrix} c_i & c_{iB} \\ 0 & c_B \end{bmatrix} + \begin{bmatrix} a_{it} & a_{Bi} \\ a_{iB,t} & a_B \end{bmatrix}\begin{bmatrix} \epsilon_{i,t-1}^2 & \epsilon_{B,t-1}\epsilon_{i,t-1} \\ \epsilon_{B,t-1}\epsilon_{i,t-1} & \epsilon_{B,t-1}^2 \end{bmatrix}$$
$$\begin{bmatrix} a_i & a_{Bi} \\ a_{iB} & a_B \end{bmatrix} + \begin{bmatrix} b_i & b_{Bi} \\ b_{iB} & b_B \end{bmatrix}\begin{bmatrix} h_{i,t-1} & h_{iB,t-1} \\ h_{iB,t-1} & h_{B,t-1} \end{bmatrix}\begin{bmatrix} b_i & b_{Bi} \\ b_{iB} & b_B \end{bmatrix}$$

A and B are symmetric matrices of parameters. C is lower triangular matrix.

A (bivariate) Student t distribution is used to allow for leptokurtic errors. The standardized probability density function is defined by

$$\epsilon_t = \begin{bmatrix} \epsilon_{it} \\ \epsilon_{Bt} \end{bmatrix} \sim t(0, H_t, v)$$

$$t(0, H_t, v) = \frac{\Gamma(n + v/2)}{\Gamma(v/2)[\phi(v-2)]^{n/2}}|H_t|^{-1/2}\left[1 + \frac{1}{v-2}\epsilon_t' H_t^{-1}\epsilon_t\right]^{-(n+v)/2} \tag{4.16}$$

where v is the degrees of freedom parameter and n(=2) is the number of equations.

To maximize this likelihood function, the most commonly used optimizer is the BHHH method which is suitable for situations in which the first order conditions contain matrices which involve derivatives that depend on past values of disturbances. The standard ARCH model can also be estimated by exact Maximum Likelihood (EML), Quasi-Maximum Likelihood (QML) and Generalized Method of Moments (GMM). The EML method refers to applying maximum likelihood estimation to equation (4.10) and (4.11) assuming a particular distribution for the error term. If the underlying distribution deviates from normality, the appropriate procedure is Quasi-Maximum Likelihood estimation. Bollerslev and Wooldridge (1992) have shown that the QML estimators are still consistent and asymptotically normal and therefore it is chosen to estimate our GARCH-M model. However, a robust standard error correction is used in this situation.

The robust standard error is valid even if the likelihood function is misspecified. Let $\theta = (\gamma_{i0}, \mu_i, \gamma_B', c_i, ..., a_i, ..., b_i)$ define the parameter vector, then an approximate variance covariance matrix for the estimator $\hat{\theta}$ is given by

$$\hat{V(\hat{\theta})} = E(\hat{\theta} - \theta_0)(\hat{\theta} - \theta_0)' \cong T^{-1}\{\hat{\vartheta}_{2D}\hat{\vartheta}_{OP}^{-1}\hat{\vartheta}_{2D}\}^{-1}$$

where

$$\hat{\vartheta}_{2D} = \left(\frac{1}{T}\right)\sum_{t=1}^{T}\frac{\partial^2 logf(Y_t;\hat{\theta})}{\partial\theta\partial\theta'} \tag{4.17}$$

$$\hat{\vartheta}_{OP} = \left(\frac{1}{T}\right)\sum_{t=1}^{T}\left[\frac{\partial logf(Y_t;\hat{\theta})}{\partial\theta}\right]\cdot\left[\frac{\partial logf(Y_t;\hat{\theta})}{\partial\theta}\right]' \tag{4.18}$$

ϑ_{2D} is the numerical approximation to the matrix of second derivatives with respect to the parameters and ϑ_{OP} is the average of the period-by-period outer products of the numerical gradient. The standard errors are simply the diagonal elements of the matrix V. In a simple Monte Carlo simulation Susmel (1994) show that tests based on these robust standard errors are conservative compared to the outer product gradient (OPG) tests. This holds for both the normal and Student t distributions. Especially for the case of leptokurtic errors, the robust test tends to have size very close to the nominal size 0.05.

Since the information matrix is not block diagonal between the parameters in the conditional mean and variance model (Bollerslev and Wooldridge (1992)), consistent estimation requires the full model to be correctly specified. To check for correct specification, a series of diagnostic tests for the variance specification is needed, such as the Ljung and Box (1978) test for autocorrelation for both the normalized residuals and squared residuals.

4.4.3 Correlated Bivariate Poisson Jump Model

Alternatively, leptokurtosis can be captured by systematic jumps. Using notations similar to that used in Chapter 3, we can modify our model as

$$
\begin{aligned}
R_{it} &= \gamma_{i0} + \mu_i \frac{\tilde{h}_{iB,t}}{h_{Bt}} (\gamma_B' x_{B,t-1}) + J_{it} + \epsilon_{it}, && (4.19) \\
R_{Bt} &= \gamma_B' x_{B,t-1} + J_{Bt} + \epsilon_{Bt}, && (4.20) \\
\epsilon_t | I_{t-1} &\sim N(0, H_t).
\end{aligned}
$$

where J_{it} is a series of jumps that is conditionally mean zero. This jump component is constructed by introducing two random variables: jump counter n_i and jump size Y_{it} as

$$J_{it} = \sum_{k=1}^{n_{it}} Y_{it,k} - (\lambda_i + \lambda_{iB})\theta_i \tag{4.21}$$

where λ_i is the independent jump intensity for the ith metal and λ_{iB} is the correlated jump intensity for the ith metal and the benchmark portfolio. The jump size is assumed be normally distributed with mean θ_i and variance δ_i^2:

$$Y_{1t,k} \sim N(\theta_1, \delta_1^2) \quad , \quad Y_{2t,j} \sim N(\theta_2, \delta_2^2) \tag{4.22}$$

The correlation between the jump sizes are constant and are denoted as ρ_{iB}.

The jump counter is controlled by a Correlated Bivariate Poisson function which gives

$$n_{it} = n_{it}^* + n_{iBt}^* \quad , \quad n_{Bt} = n_{Bt}^* + n_{iBt}^* \tag{4.23}$$

and ρ_{iB} is the constant correlation between the jump sizes. Both the metal price and benchmark portfolio may experience systematic jumps and the jumps can be separated into independent and correlated jumps. Independent jumps are initiated by the n_{it}^* for the metal price and n_{Bt}^* for the benchmark portfolio, whereas correlated jumps are governed by n_{iBt}^*.

In the presence of jumps, the time varying risk premia will be not only composed of the systematic risk associated with the benchmark portfolio, but also of the risk associated with jumps. The covariance conditional on k jumps arriving to the excess returns and j jumps to the benchmark is given by

$$Cov(R_t, R_{Bt}) = \tilde{h}_{iB} = h_{iB} + \rho_{iB}\sqrt{kj}\delta_1\delta_2 \tag{4.24}$$

and the conditional variance is

$$Var_{t-1}(R_{Bt}) = \tilde{h}_B = h_B + j\delta_2^2. \tag{4.25}$$

Therefore, the original model can be rewritten as

$$\begin{aligned} R_{it} &= \gamma_{i0} + \mu_i \frac{\tilde{h}_{iB,t} + \rho_{iB}\sqrt{(\lambda_i + \lambda_{iB})(\lambda_B + \lambda_{iB})}\delta_1\delta_2}{\tilde{h}_{Bt}}(\gamma_B' x_{B,t-1}) \\ &\quad + J_{it} + \epsilon_{it}, && (4.26) \\ R_{Bt} &= \gamma_B' x_{B,t-1} + J_{Bt} + \epsilon_{Bt}, && (4.27) \end{aligned}$$

Note that the price of risk μ_i reflects the risk premia associated with the benchmark and systematic jumps. If the correlation between the jump sizes is ρ_{iB}=0, the risk premium term will collapse to the original form in Equation (4.14). The log likelihood function takes the form of Equation (3.38).

4.5 Data

The resource data correspond to four metals currently traded on the LME: copper, lead, sliver, and zinc. Monthly price data are for the period January 1955 through December 1990. These data are obtained from the American Bureau of Metal Statistics. The interest rates used are the US 30-Day Treasury Bill rates from the Center for Research in Security Prices (CRSP) database. The monthly average operating cost data are extrapolated from Young and Ryan's (1996) data which are derived from the mining data provided by Canadian Minerals Yearbook and The General Review of the Mining Industry. The operating cost data include costs associated with wages, materials, supplies, fuel, and energy.

Table 4.1 provides summary statistics for the four metals. The average rate of return for the gross price data is around 0.23% monthly over the 36 year span of data. The rates of return are spread evenly across positive and negative values, and a common decreasing trend around 1980's for some prices is found. The decreasing trend has been a major problem in the primary metal industry in last two decades and it is evident in the lower panel of Table 4.1 where lead and zinc have negative average returns in 1980's. This is consistent with the findings of high volatility in Slade (1982) and Slade (1991). In Slade's (1991) paper, she concludes that the high volatilities of metal prices can be explained by changes in market structure. Although her data set covers only the period of January 1970 to December 1986, a similar pattern is found in our data set.

The raw metal prices are graphed in Figure 4.1 where each metal price is measured in U.S. cents per pound, except for the price of silver which is U.S. cents per troy ounce. Most metals experience high fluctuations over the sample period, except silver which varies predominantly within a few abnormal episodes. Figure 4.2 shows the rates of return on prices for these metals along with the Treasury Bills and the CRSP index. A different view of silver is depicted in panel (c) as opposed to the previous figure. Silver returns do exhibit high variations after around 1967 and most of these movements may be mistakenly viewed as small relatively to the big jumps in the early 1980's.

A simple check for normality shows that the skewness coefficients are very close to zero, whereas the kurtosis coefficients vary from 2.5473 to 21.9921. This indicates that a more general distribution should be used to model these rates of return. The kurtosis coefficient for silver is extraordinarily large at 21.9921, which could be due to speculation in the early 1980's, spurred by activities of the Hunt brothers.

The CRSP index is chosen as the benchmark portfolio. The CRSP index is a value-weighted index of the US equity market including all distributions. Since the time-varying risk premium is a function of the correlation between the benchmark portfolio and the resource return, the correlation coefficients between the rates of return of each metal and the various benchmark portfolio are very important to the model. The correlation coefficients for the CRSP index is as high as 0.29 for copper to as low as 0.01 for zinc.

The average cost data are depicted in Figure 4.3 as percentage of the raw metal prices. The proportion of average cost changes substantially over these 36 years and in some cases the zinc market actually has average production cost that are above the nominal price level. The average cost remains a small portion of the silver price, whereas the average cost for lead varies from as low as 22% in 1955 to as high as 95% in 1985.

4.6 Results

Before examining the risk premia, we will first discuss the general performance of each model in terms of the GARCH structure, the use of the Student t distribution, the incorporation of systematic jumps, and difference associated with the use of gross versus net metal prices. A detailed comparison of the time-varying risk premia will be presented later in this section.

4.6.1 Gross Metal Prices

Table 4.2 reports the bivariate GARCH-M results based on gross metal price returns. Note that the Ljung Box statistics show that an AR(1) term must be added to the metal price equation to capture all of the serial correlations. The use of the bivariate model is supported by the significant a_{Bi} and b_{Bi} estimates which capture the cross equation effect. If these coefficients are zero, the conditional variances will be a function of only their own lagged values and the lagged error squares. In this case, the advantage of using the bivariate model would be undermined by not using the information on conditional covariances in the variance-covariance structure.

The risk premium parameters, μ, are large, but insignificant in all cases. However, it is interesting to note that all metals report a negative risk premium parameter. The diagnostic statistics in Table 4.3 shows that the standard BEKK-GARCH-M model fits the data reasonably well and no uncaptured serial correlations can be found in any metals with the exception of the squared standardized residuals from lead and silver.

Another interesting feature of these results are the coefficients for the benchmark portfolio. Since the same benchmark portfolio is used for all four metals, a similar set of values are expected for the parameters related to the benchmark portfolio alone. For example, the parameters, a_B and b_B, have similar size in all cases and the sign difference is of no importance because of the quadratic form. Similarly, the constant term γ_{B0} and the AR(1) term γ_B have values around the same range across metals. In addition, the constant terms are always significant, whereas the AR(1) terms are always insignificant. Consistency in the estimated parameters across metals is found with the Student t distribution.

A noticeable difference in silver, as compared to other metals, is found in the relatively large estimated coefficients a_{Bi} and b_{Bi}. It is not surprising that silver behaves differently from the other metals given the price surge in the early 1980's. One anomaly is the large kurtosis and skewness coefficients for silver, which casts doubts on the usefulness of the t distribution for this metal. Since silver is a precious metal it may be that a precious metals commodity index might be more appropriate as the benchmark portfolio. Alternatively, a more general distribution allowing for skewness might be required for silver. However, as noted above, we rely on the robust standard errors to deal with any deviation from normality.

The London Metal Exchange underwent a basic restructuring in 1987. This involved changing from principal market to a clearing house market, the introduction of margins and formal options. Dummy variables are introduced into the model to test for the effects of restructuring on the stability of the parameters. Two types of dummy variables are added to test for the presence of a structural break. The constant dummy tests the shift of the constant term in the model, whereas the multiplicative dummy checks the stability of the coefficients. The bivariate models were re-estimated and no structural breaks were found in the empirical models.

4.6.2 Net Metal Prices

The results which take into account the cost data are reported in Table 4.4. Note that with these net return series the AR(1) term in the metal equations is no longer required to remove all the serial correlations in the standardized residuals of the metal equation. The degrees of freedom parameter once again is significant in all cases, although the size of these parameters are relatively larger than the ones with gross prices, implying leptokurtic distributions. The GARCH parameters reveal similar patterns for the four metals and the constant terms for the benchmark portfolio are very consistent.

A significant negative time-varying risk premium parameter is discovered for copper price with the incorporation of production costs. Recall that a negative price parameter does not imply negative risk premium. This depends on the sign of the covariance term between the metal returns and the benchmark portfolio. The same measure for the other three metals are found to be insignificant regardless of using net or gross metal prices to construct the return series. The diagnostic statistics in Table 4.5 reveal the adequacy of using the bivariate GARCH-M model to capture the dynamics. The Ljung Box Statistic on the squared standardized residuals reveals serial correlations from the silver price and the same result was found in the case of gross prices.

Taking into account of the production cost does have implications for the estimated parameters. For example, the benefit of estimating a bivariate model is less apparent as cross equation GARCH parameters become small and insignificant. The other interesting result is the change of sign and magnitude of the risk premium parameters from negative to positive for 3 of the 4 metals. The dramatic change of parameter estimates is a direct consequence of high variations in average cost data as depicted in Figure 4.3.

4.6.3 Net Metal Prices with Conditional Jump Dynamics

Alternatively, the leptokurtic distribution can be modeled by the correlated bivariate Poisson jump model as discussed earlier and the results for this CBP-GARCH-M model are presented in Table 4.6. Note that uncaptured serial dependence is displayed by the Ljung Box statistics on the squared standardized residuals for lead and zinc.

In general, allowing for jumps in the model does not provide significant improvement on either the likelihood values or the diagnostic statistics in copper, lead, and zinc.

Jump parameters are mostly insignificant and correlated jumps can be rejected at 5% level for all metals. In other words, there are no simultaneous shocks to both the metal price and the benchmark portfolio. For copper, lead, and zinc, the independent jump intensity parameter λ_B for the benchmark is always insignificant. A significant jump intensity λ_i is found in the copper equation. However, the likelihood value has shown that the bivariate jump model does no better than a simple bivariate Student t distribution.

A noticeable improvement in the case of silver can be found in the likelihood value, the Ljung Box statistics, and the significance of the jump parameters. The likelihood value is the best overall as compared to the previous two specifications. The CBP-GARCH-M model is the only one model that can reject serial correlations in all cases at 5% level. The results suggest that silver experiences many abrupt changes over the last 36 years with jump sizes around -1% to 1% being the most frequent.

Turning to the time-varying risk premia, the estimated price coefficients are insignificant for copper, lead, and zinc, but significant and positive for silver. We have are no plausible explanations for this finding. First, a significant risk premium can only be found by fully capturing all the dynamics in both the metal and benchmark series and the CBP-GARCH-M model performs the best. Second, although evidence suggests that the jump model fits the silver series well and a time varying risk premium is found, we have to be cautious about this finding as the significant jump component may merely be capturing the Hunt's brother incident in the early 1980's, creating such a large jump in the return series.

4.6.4 Discussion

The possibility of time varying risk premia in non-renewable resource markets is an interesting question. However, it is difficult to identify risk premia given the constraint of estimating a correctly specified model. We find weak evidence of time varying risk premia in two metals: copper and silver. The metals prices being studied in this chapter are well represented by the bivariate GARCH model with or without augmenting the jump dynamics.

A significant negative price parameter for the time varying risk premium in net copper price is revealed with the GARCH-M model with Student t distribution. The implied risk premia are depicted in panel (c) of Figure 4.4 along with the return series on the top panel (a). In general, the series experiences little variation before 1981 and only a few episodes generating large negative values for the risk premium in the early and late 1980s. Positive risk premia are infrequent and mostly less than 1%. It is of interest to note that it is only after the change in the market structure for copper (See Slade (1988)) that variations in risk premia over time for copper are observed. In general, there is no strong evidence supporting time varying risk premia as an important part of copper returns for most of the time horizon considered.

The time varying risk premia for silver are graphed in panel (d) of Figure 4.4 using the significant price parameter μ. The most important observation is the magnitude of these risk premia within zero to one percent for the most part with no substantial variations recorded until the period of the early 1980's. These results are consistent with the findings in Young and Ryan (1996) where the risk premia in copper and silver are significant, however, with a very small size.

In summary, time-varying risk premia can be detected in two of the four metal prices being studied depending on the modeling strategy. The two models supply different results suggesting the appropriateness of incorporating the risk premia in empirical models. The presence of time-varying risk premia can be tested by using a benchmark portfolio and metal returns. These results not only provide cautionary notes on examining the Hotelling rule with non-renewable resource returns, but also extend the current empirical literature on risk premia (Slade and Thille (1995) and Young and Ryan (1996)).

4.7 Conclusions

In this chapter a bivariate GARCH-M model is used to incorporate time-varying risk premia into a simple Hotelling model. This approach is based on the transformation of the risk-adjusted Hotelling rule into the consumption-based Intertemporal Asset Pricing Model, where the time-varying risk premia are modeled by the ratio of the conditional covariance between metals and a benchmark portfolio and the conditional variance of the benchmark.

The presence of system jumps (market crashes) and its effect on risk premia are also examined with the same bivariate model augmented with Correlated Bivariate Poisson (CBP) jumps. Using four metal prices (copper, lead, silver, and zinc) covering four decades, we find that the time-varying risk premium is most often not an important determinant of the nonrenewable resource return regardless of the modeling strategy.

The results reveal the consequences of mis-specification in the context of searching for time varying risk premia. Misleading results may come from neglecting marginal cost data, inadequately capturing of volatility dynamics, and failure to recognize the presence of systematic jumps.

4.8 Appendix B

4.8.1 Derivation of Equation (4.7) and (4.8)

Dividing through by $U'(C_1)$ in equations (4.2) to (4.4), we have

$$-1 + EM_s \cdot [1+\rho] = 0 \tag{4.28}$$

$$-1 + EM_s \cdot [1+r] = 0 \tag{4.29}$$

$$1 \cdot (P_1) - EM_s \cdot (P_2) = 0 \tag{4.30}$$

$$M_s = \beta(U'(C_2)/U'(C_1)) \tag{4.31}$$

Rearranging gives:

$$EM_s r = 1 - EM_s \tag{4.32}$$

$$1 - EM_s = EM_s \frac{\Delta P}{P} \tag{4.33}$$

Substituting equation (4.32) into (4.33) yields:

$$EM_s r = EM_s \frac{\Delta P}{P} \tag{4.34}$$

If we assume that the marginal rate of substitution M_s is constant, Equation (4.34) becomes

$$E(r) = E\left(\frac{\Delta P}{P}\right)$$

The expected rate of increase of the net price is equal to the expected rate of return. Assuming that M_s is not a constant, the right hand side of equation (4.34) can be rewritten as

$$EM_s \frac{\Delta P}{P} = EM_s E \frac{\Delta P}{P} + Cov\left(M_s, \frac{\Delta P}{P}\right)$$

Similarly, the right hand side becomes

$$EM_s r = EM_s E(r) + Cov(M_s, r)$$

Substituting these into Equation (4.34) gives

$$EM_sE\frac{\Delta P}{P} + Cov\left(M_s, \frac{\Delta P}{P}\right) = EM_sE(r) + Cov(M_s, r) \tag{4.35}$$

$$E\frac{\Delta P}{P} = E(r) - \frac{Cov(M_s, \Delta P/P) - Cov(M_s, r)}{E(M_s)} \tag{4.36}$$

$$E\frac{\Delta P}{P} = E(r) - \frac{Cov(M_s, \Delta P/P - r)}{E(M_s)} \tag{4.37}$$

Equation (4.8) can be derived using similar steps, utilizing the following two formulas from Equation (4.28) and (4.30)

$$\begin{aligned} \rho EM_s &= 1 - EM_s \\ P_1 - EM_sP_2 &= 0 \end{aligned}$$

Table 4.1: Summary Statistics (Rates of Return 1955-1990)

Gross Metal Return Data						
	Copper	Lead	Silver	Zinc	T-Bill	CRSP
Mean	0.2264	0.1829	0.1362	0.3817	0.4747	0.4487
S.D.	6.8874	6.0483	8.4846	5.7389	0.2533	4.3287
Skewness	-0.3379	0.2852	-0.5735	0.2593	1.0209	-0.3886
Kurtosis	2.5473	3.2229	21.9921	4.9337	1.3228	2.0797
Average Over Each Decade						
	Copper	Lead	Zinc	Silver	T-Bill	CRSP
1950's	-0.4954	-0.6028	0.1468	0.1162	0.1875	0.9907
1960's	0.7332	0.4247	0.5465	0.1385	0.3147	0.4154
1970's	0.2116	1.0504	1.2920	0.7362	0.5123	0.1044
1980's	0.0875	-0.4801	-1.2900	0.3971	0.7090	0.5539
Net Metal Return Data						
	Copper	Lead	Silver	Zinc		
Mean	0.1414	0.2943	0.1356	0.7201		
S.D.	10.8373	13.5865	8.4867	11.1358		
Skewness	-0.5247	-1.2275	-0.5702	4.5828		
Kurtosis	46.0336	62.0524	21.9374	79.4051		

Figure 4.1: Metal Prices (1955 - 1990)

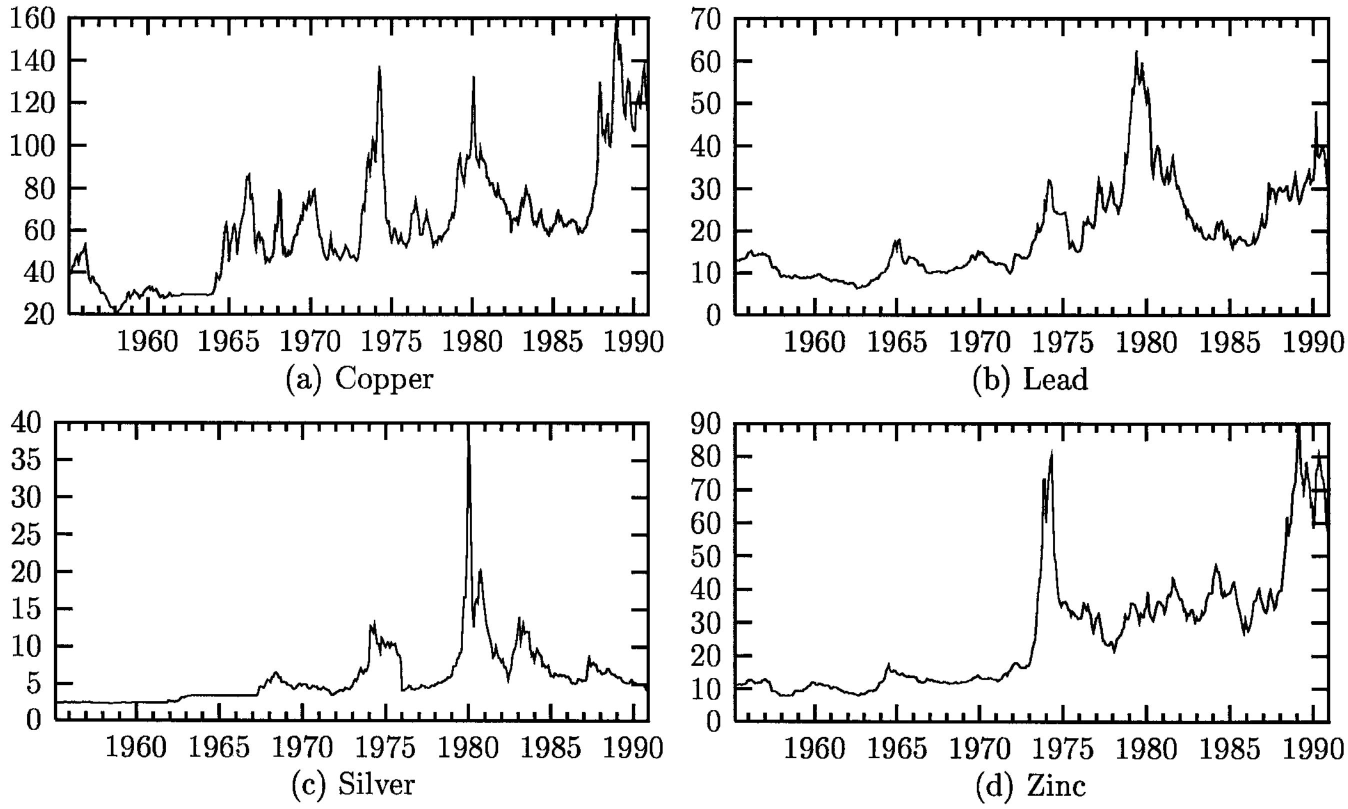


Table 4.2: Bivariate GARCH-M model with t distribution (Gross Metal Prices and CRSP Monthly Data)

Parameter	Copper	Lead	Silver	Zinc
γ_{i0}	**-0.0018	*-0.0047	**-0.0020	-0.0018
	(0.0005)	(0.0028)	(0.0003)	(0.0024)
γ_0	**0.2250	**0.3041	**0.3196	**0.2627
	(0.0537)	(0.0535)	(0.0574)	(0.0488)
γ_{B0}	**0.4121	**0.5436	**0.4448	**0.5311
	(0.2209)	(0.1872)	(0.1851)	(0.2148)
γ_B	-0.0535	0.0446	0.0417	-0.0081
	(0.0254)	(0.0391)	(0.0477)	(0.0336)
c_i	9.90E-08	**0.0001	1.30E-08	0.00002
	(0.0015)	(0.0153)	(0.0004)	(0.0065)
c_{iB}	0.0005	**-0.0190	0.0001	**0.0058
	(0.0004)	(0.0044)	(0.0006)	(0.0024)
c_B	**3.0524	**1.8525	**2.8532	**1.2786
	(0.7605)	(0.5794)	(0.3900)	(0.3406)
a_i	**0.7064	**0.6272	**0.8402	**0.4397
	(0.0593)	(0.0919)	(0.0656)	(0.0605)
a_{iB}	0.0001	0.0004	**0.0011	**-0.0010
	(0.0001)	(0.0008)	(0.0001)	(0.0004)
a_{Bi}	**-7.8877	-4.8464	**-17.3967	**-7.1399
	(3.0692)	(3.9676)	(6.0760)	(2.7942)
a_B	**0.1482	*-0.1287	-0.1146	**-0.1451
	(0.0492	(0.0659)	(0.0776)	(0.0670)
b_i	**0.8175	**0.7376	**0.7643	**0.8989
	(0.0208)	(0.0832)	(0.0216)	(0.0242)
b_{iB}	-0.0001	**0.0015	-0.00001	**-0.0009
	(0.0001)	(0.0006)	(0.0001)	(0.0001)
b_{Bi}	**6.7399	**6.3325	**10.2094	**6.2986
	(2.4446)	(3.1111)	(2.2450)	(1.7060)
b_B	**0.6895	**0.8930	**0.6456	**0.9333
	(0.1704)	(0.0659)	(0.1140)	(0.0281)
$1/v$	**0.1626	**0.1333	**0.0300	**0.1045
	(0.0311)	(0.0324)	(0.0072)	(0.0290)
μ	-8.4837	-4.8796	-0.2358	-6.8407
	(5.3761)	(3.7367)	(1.3626)	(4.4428)
log likelihood	-578.01	-553.64	-458.67	-494.89

Standard errors are in parenthesis. * refers to 10% level of significance and ** refers to 5% level of significance.

Table 4.3: Diagnostic Statistics Bivariate GARCH-M model with t distribution (Gross Metal Prices and CRSP Monthly Data)

	Copper	Lead	Silver	Zinc
Q_i	10.81	6.38	11.92	10.62
	[0.37]	[0.78]	[0.29]	[0.38]
Q_i^2	10.85	19.53	20.59	15.74
	[0.36]	[0.03]	[0.02]	[0.10]
Q_B	16.04	11.09	11.01	15.87
	[0.09]	[0.34]	[0.35]	[0.10]
Q_B^2	13.49	9.43	10.29	4.52
	[0.19]	[0.49]	[0.41]	[0.92]
Q_{iB}	16.17	17.72	8.05	11.72
	[0.09]	[0.05]	[0.62]	[0.30]

Standard errors are in parenthesis. p-values are in square brackets. Q^2 is the modified Ljung-Box portmanteau test, robust to heteroskedasticity, for serial correlation in the squared standardized residuals with 10 lags for the respective models.

Table 4.4: Bivariate GARCH-M model with t distribution (Net Metal Prices and CRSP Monthly Data)

Parameter	Copper	Lead	Silver	Zinc
γ_{i0}	**-0.0030	-0.0010	0.0015	-0.0043
	(0.0006)	(0.0041)	(0.0037)	(0.0042)
γ_{B0}	**0.7350	**0.6563	**0.7788	**0.5996
	(0.2091)	(0.1988)	(0.1939)	(0.2016)
γ_B	-0.0350	-0.0177	0.0168	0.0335
	(0.0347)	(0.0322)	(0.0439)	(0.0531)
c_i	3.08E-06	0.0002	0.0017	0.0003
	(0.0149)	(0.0227)	(0.1459)	(0.0039)
c_{iB}	**-0.0021	**0.0213	**0.0830	**-0.0277
	(0.0010)	(0.0061)	(0.0334)	(0.0064)
c_B	**1.4456	**1.619	3.8880	**2.1850
	(0.6979)	(0.5507)	(2.3127)	(0.6229)
a_i	**1.2843	**0.8378	**1.5980	**0.4388
	(0.1752)	(0.1543)	(0.6024)	(0.0859)
a_{iB}	-0.0001	0.0011	-0.0031	-0.0014
	(0.0004)	(0.0015)	(0.0050)	(0.0016)
a_{Bi}	0.3174	0.4030	0.4578	1.4451
	(0.6463)	(0.7670)	(0.8594)	(1.2110)
a_B	**0.1868	*-0.1371	-0.2494	**0.2774
	(0.0690)	(0.0731)	(0.1941)	(0.0944)
b_i	**0.6608	**0.7221	**-0.4991	**0.8691
	(0.0395)	(0.0847)	(0.0804)	(0.0305)
b_{iB}	0.0001	**-0.0017	**0.0056	**0.0024
	(0.0001)	(0.0007)	(0.0020)	(0.0005)
b_{Bi}	-0.1615	0.3348	-0.1854	-0.6550
	(0.1785)	(0.5629)	(0.5602)	(0.5771)
b_B	**0.9550	**0.9399	**-0.8812	**0.9031
	(0.0345)	(0.0383)	(0.0854)	(0.0424)
$1/v$	**0.3310	**0.2788	**0.4335	**0.3428
	(0.0459)	(0.0402)	(0.0579)	(0.0409)
μ	**-2.8330	6.0109	2.8985	0.0103
	(1.5325)	(3.8117)	(2.2067)	(1.9369)
log likelihood	-1044.81	-801.45	-1039.21	-856.85

Standard errors are in parenthesis. * refers to 10% level of significance and ** refers to 5% level of significance.

Table 4.5: Diagnostic Statistics Bivariate GARCH-M model with t distribution (CRSP Monthly Data)

	Copper	Lead	Silver	Zinc
Q_i	18.63	15.38	9.36	10.04
	[0.05]	[0.11]	[0.49]	[0.43]
Q_i^2	9.83	16.29	29.60	15.29
	[0.45]	[0.09]	[0.00]	[0.12]
Q_B	15.16	13.93	12.00	11.58
	[0.12]	[0.17]	[0.28]	[0.31]
Q_B^2	11.36	11.53	11.67	6.02
	[0.32]	[0.31]	[0.30]	[0.81]
Q_{iB}	8.50	13.14	18.16	6.01
	[0.57]	[0.21]	[0.05]	[0.81]

Standard errors are in parenthesis. p-values are in square brackets. Q^2 is the modified Ljung-Box portmanteau test, robust to heteroskedasticity, for serial correlation in the squared standardized residuals with 10 lags for the respective models.

Table 4.6: Estimates of CBP-GARCH-M model with normal distribution (Risk Premia and Jump Components)

Parameter	Copper	Lead	Silver	Zinc
μ	-1.4233	-8.4632	**0.6907	1.3149
	(1.6070)	(5.4123)	(0.2851)	(2.3517)
λ_i	**0.0304	0.0082	**0.0811	0.0350
	(0.0109)	(0.0274)	(0.0227)	(0.0298)
δ_i	**7.5123	**0.5380	**1.3320	**0.6339
	(1.4767)	(0.1512)	(0.2652)	(0.1081)
θ_i	-0.7960	0.0121	-0.0300	0.0314
	(1.9002)	(0.1216)	(0.2274)	(0.1127)
λ_B	0.0410	0.1057	**0.1777	0.0361
	(0.0426)	(0.1010)	(0.0571)	(0.0478)
δ_B	0.0036	**6.2431	7.40E-11	**6.7298
	(0.1262)	(1.8461)	(0.0766)	(1.8921)
θ_B	**-8.4471	-0.2820	**-5.9625	-2.4352
	(2.4700)	(1.1355)	(0.8377)	(2.0786)
λ_{iB}	0.0078	0.0457	0.0104	*0.0469
	(0.0079)	(0.0287)	(0.0140)	(0.0263)
ρ	-0.9999	-0.0595	-0.2417	-0.0889
	(27.7284)	(0.2819)	(177.0260)	(0.1188)
log likelihood	-1099.85	-801.79	-1031.25	-845.13
Q_i	13.13	10.75	8.66	9.30
	[0.21]	[0.37]	[0.54]	[0.50]
Q_i^2	6.50	20.40	17.69	18.78
	[0.77]	[0.03]	[0.06]	[0.04]
Q_B	10.85	10.88	13.30	11.50
	[0.36]	[0.36]	[0.20]	[0.31]
Q_B^2	8.41	12.57	8.10	6.62
	[0.58]	[0.24]	[0.61]	[0.75]
Q_{iB}	7.98	9.25	16.78	6.20
	[0.62]	[0.50]	[0.07]	[0.79]

Standard errors are in parenthesis. p-values are in square brackets. Q^2 is the modified Ljung-Box portmanteau test, robust to heteroskedasticity, for serial correlation in the squared standardized residuals with 10 lags for the respective models. * refers to 10% level of significance and ** refers to 5% level of significance.

Table 4.7: Estimates of CBP-GARCH-M model with normal distribution (GARCH Parameters)

Parameter	Copper	Lead	Silver	Zinc
γ_{i0}	-0.0286	-0.0054	-0.0044	-0.0021
	(0.0731)	(0.0078)	(0.0211)	(0.0097)
γ_{B0}	0.3235	**0.3935	*0.3681	**0.5116
	(0.2145)	(0.1940)	(0.1977)	(0.2139)
γ_B	0.0576	**0.0634	0.0012	0.0344
	(0.0486)	(0.0323)	(0.0419)	(0.0580)
c_i	-0.0017	0.0000	0.0070	**-0.0226
	(0.0885)	(0.0490)	(0.0649)	(0.0031)
c_{iB}	**-0.0412	-0.0052	**-0.0512	0.0011
	(0.0085)	(0.0111)	(0.0095)	(0.0394)
c_B	**-0.9669	**3.5101	*0.3777	**2.0816
	(0.4183)	(0.3058)	(0.2134)	(0.8152)
a_i	**0.7237	**-0.6951	**0.9543	**0.3290
	(0.0808)	(0.0706)	(0.0770)	(0.0526)
a_{iB}	-0.0021	0.0006	-0.0017	-0.0013
	(0.0014)	(0.00007)	(0.0016)	(0.0010)
a_{Bi}	0.0022	0.0294	**0.5694	-0.0231
	(0.1427)	(0.6720)	(0.0016)	(0.7757)
a_B	**-0.1753	**0.1203	**-0.2365	**-0.2354
	(0.0471)	(0.0583)	(0.0353)	(0.0672)
b_i	**0.5904	**0.6251	0.0034	**-0.8146
	(0.0674)	(0.0542)	(0.0150)	(0.0368)
b_{iB}	0.0044	**-0.0057	-0.0013	0.0009
	(0.0023)	(0.0012)	(0.0017)	(0.0068)
b_{Bi}	-0.3115	-1.0685	**0.9056	-0.5080
	(0.3196)	(1.5583)	(0.0354)	(2.8257)
b_B	**-0.9481	0.1008	**-0.9437	**0.7920
	(0.0285)	(0.5069)	(0.0146)	(0.1634)
log likelihood	-1099.85	-801.79	-1031.25	-845.13

Standard errors are in parenthesis. * refers to 10% level of significance and ** refers to 5% level of significance.

Figure 4.2: Rate of Returns (1955 - 1990)

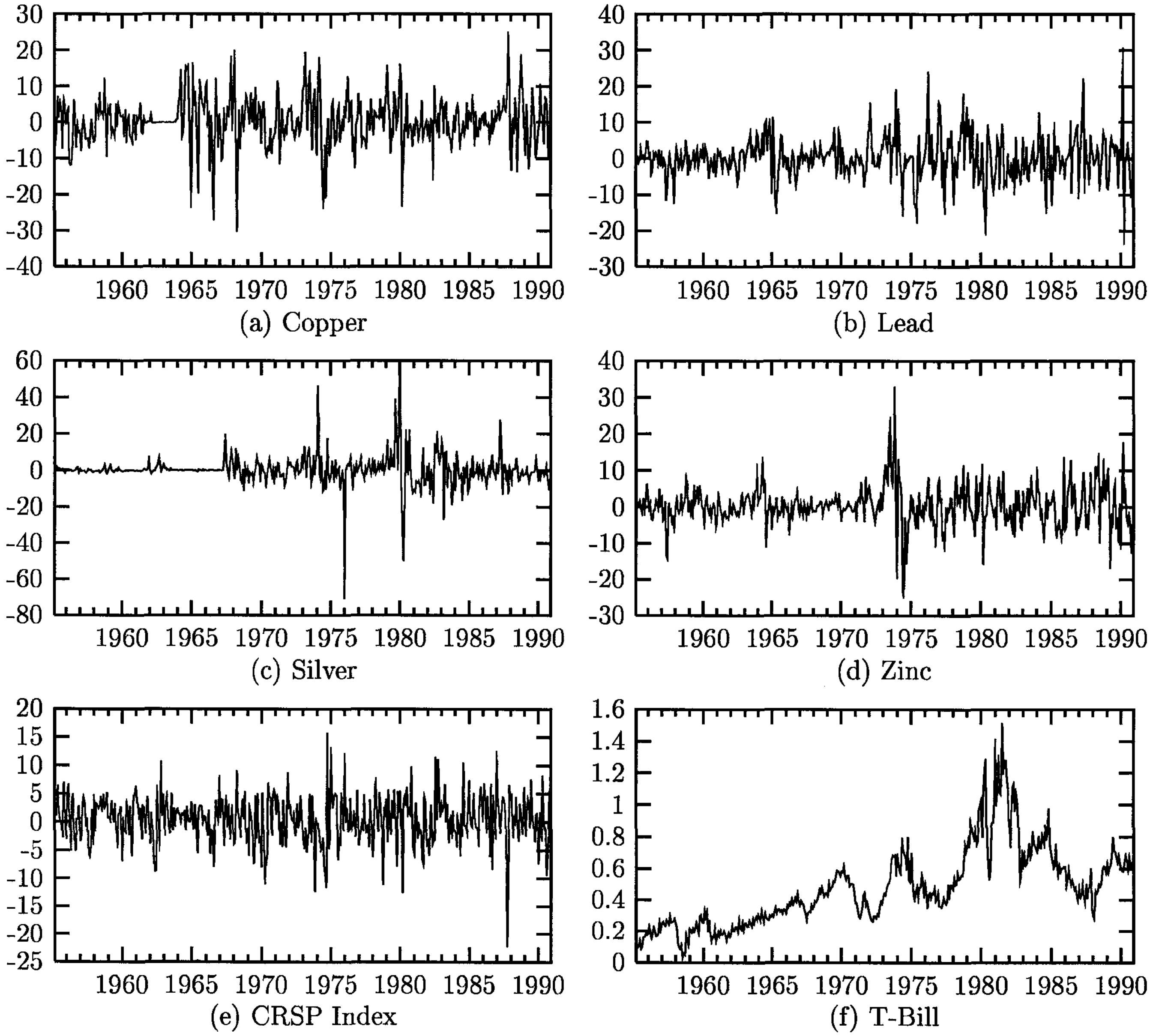
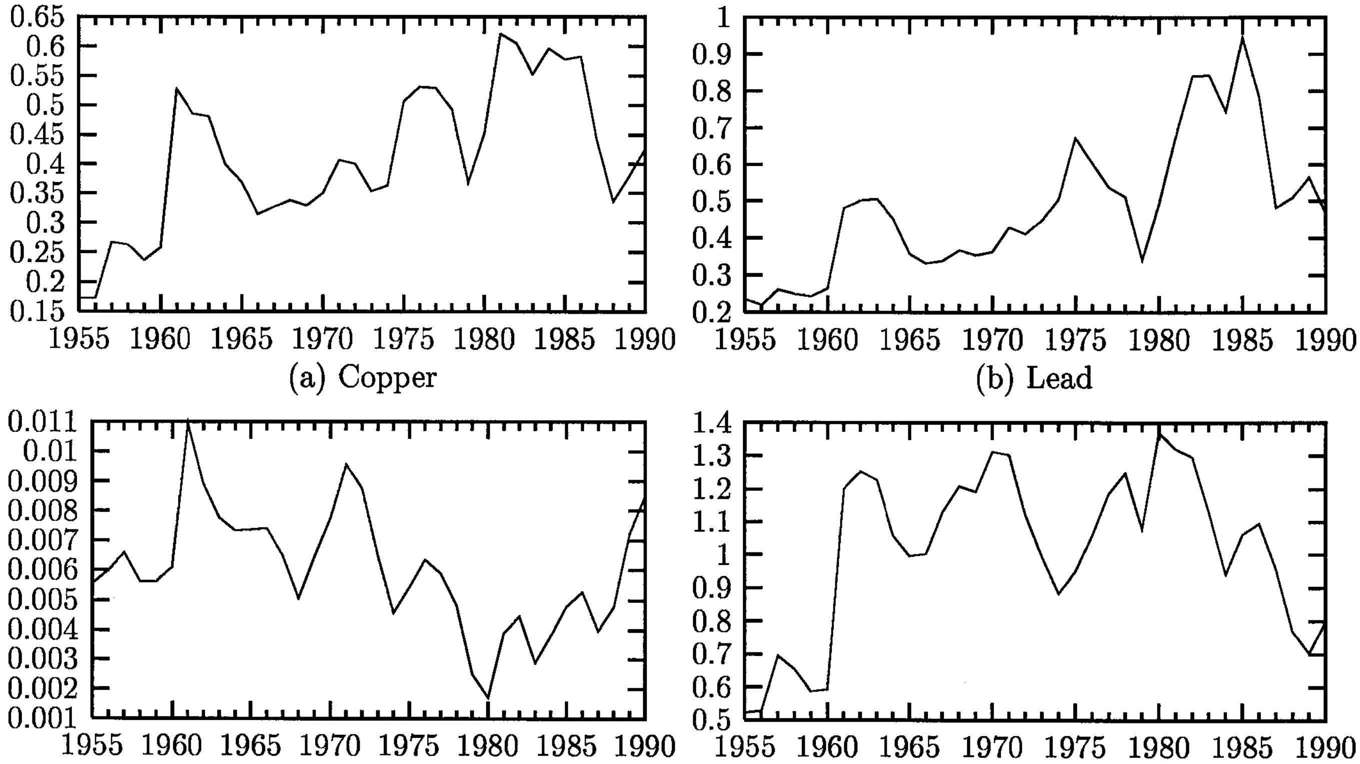


Figure 4.3: Average Cost as Percentage of Metal Price (1955 - 1990)



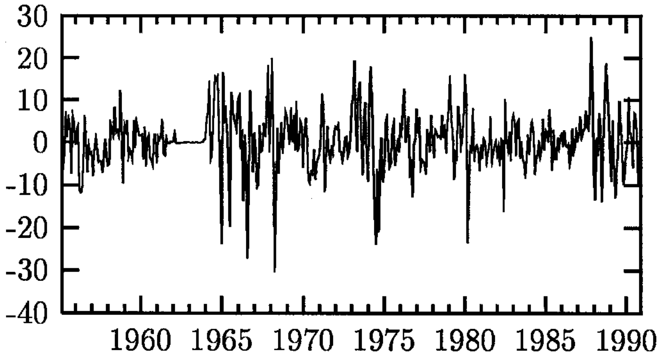
(a) Copper

(b) Lead

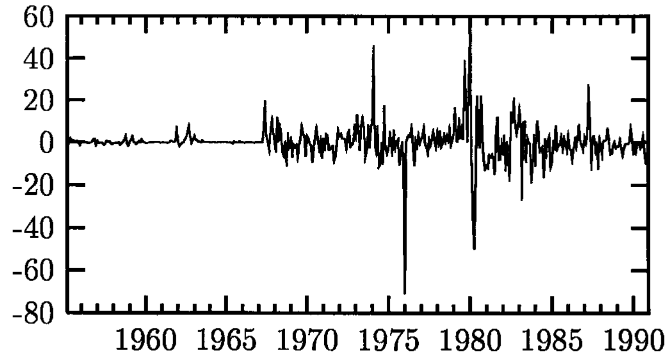
(c) Silver

(d) Zinc

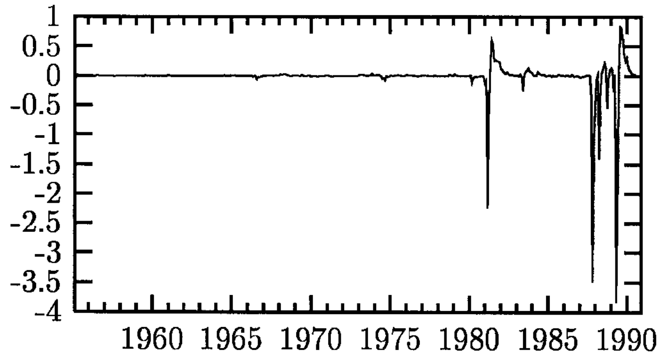
Figure 4.4: Time Varying Risk Premia



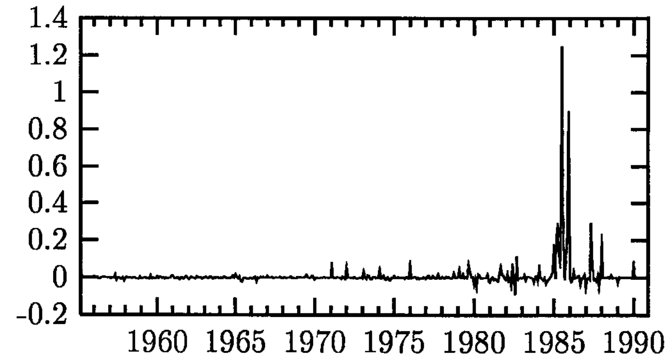
(a) Rates of Return (Copper)



(b) Rates of Return (Silver)



(c) Risk Premium (Copper)



(d) Risk Premium (Silver)

Chapter 5

Conclusion

5.1 Summary

This thesis provides an avenue for further development in nonlinear modeling of financial asset returns. A statistical jump model is developed to better capture unusually large movements in financial series ranging from stock markets to foreign exchange markets, and commodity markets.

In Chapter 2, we show that changes in returns of unusually large magnitude which occur in response to the entry of unanticipated news into the stock market can be successfully modeled in terms of a Poisson jump model with a time varying jump frequency. Furthermore, the sizes of market crashes are found to be closely related to market conditions in the previous period. The empirical results indicate that, for example, a rally is most likely to occur after a severe crash in the stock market. A new empirical insight into market dynamics is obtained by comparing the jump frequency and market movements. One-step ahead forecasts of jump frequency may serve as good indicators of severe market crashes.

The Correlated Bivariate Poisson jump model introduced in Chapter 3 identifies significant simultaneous shocks in both the Canadian dollar and Japanese Yen against U.S. dollar exchange rates. In addition, shocks specific to a single currency also become evident through appropriate statistical modeling. This finding provides new insight into the comovement of the two series. Furthermore, it allows for the construction of more accurate volatility forecasts. We have shown that the restriction of a constant theoretical upper bound of jump counter correlation is eliminated by the use of an approximate ARMA intensity structure. The data reject the hypotheses of no correlated jumps and of constant jump frequency. We also establish a relationship between correlated jump frequency and correlations across returns. We find that as simultaneous shocks intensify, the correlations between returns tend to have a positive sign once the number of jumps exceed certain threshold.

Time-varying risk premia in metal prices provide the main focus of Chapter 4. A bivariate GARCH-M model augmented with correlated Poisson jumps is employed for the analysis. Of the four metals under study (copper, lead, silver, and zinc), we find weak evidence of risk premia in copper and silver returns, although with very small magnitude. The results show signs of remaining misspecification, including uncaptured serial correlations and inadequately modeled excess kurtosis in the residuals. The consequences of this misspecification may include misleading results regarding time-varying risk premia.

To answer the questions posed in the introduction, we find evidence that there is a systematic pattern in the arrival rate of unusually large market movements and modeling this information appropriately provides a better empirical understanding of financial market movement. The magnitude of a crash will certainly affect the characteristics of future crashes. According to our results, the sizes of future jumps/crashes are directly related to the market's history as captured within the rates of return. The likelihood of crashes does change over time, depending on the state of the market. Furthermore, unanticipated news leading to these crashes should be classified into sources that generate a large movement in a single series and into sources generating simultaneous large changes in multiple series. Appropriate classification allows for increased accuracy of inferences for various types of future market crashes.

As a cautionary note, each of the first two chapters has proposed a new statistical model to capture systematic patterns in asset returns. These empirical jump models work well in variety of situations; however, the underlying source of these jumps remains unknown. A theoretical model has yet to be developed to reveal the driving force behind these statistical features. This is a drawback of my time series analysis. On the other hand, applying a structural model to test an economic theory as in Chapter Four fails to find significant jump components possibly because the theoretical structure may be too restrictive for modeling real world metal prices. Researchers should be reminded that one may only improve the model performance by incorporating additional information with a structure that is supported by sound economic theory.

5.2 Future Work

There are many potential future avenues of research that flow from the research undertaken in this thesis. Future research may focus on the development of the techniques and modeling strategies proposed or on particular applications.

The proposed modeling technique is suitable for a wide variety of situations. Any data series that is subject to frequent or infrequent shocks can be examined by the Poisson jump model. One extension is to use an alternative distribution to replace the Poisson function in constructing the Autoregressive Jump Intensity. A major weakness of using the Poisson distribution is the restriction that the mean and variance must be equal. This implies that as the number of jumps increases, the variance of the jump counter also increases. The negative binomial distribution is one possible candidate for relaxing this assumption.

For the multivariate framework, it would also be worthwhile to investigate different forms of the multivariate Poisson distribution. One possibility is to introduce nonlinear combinations of independent and correlated Poisson variables to construct a multivariate Poisson distribution. It may lead to other possible solutions to the theoretical upper bound imposed by the multivariate Poisson distribution which was used in this thesis. Developing a flexible multivariate Poisson distribution may contribute to areas of econometrics which currently rely on multiple discrete counting processes.

The GARCH-M model used in Chapter 4 can be further extended to incorporate measurable shocks as risk premia. The potential gain in measuring the monetary values of these shocks is large. There is no reason to believe that risk premia are uniquely driven by the conditional heteroskedastic variance and not the measurable shock. Market participants are likely to demand a risk premium once they learn of the existence of such a systematic patterns.

One potential application of the general framework expounded in this thesis is a closed-form option price formula given the ARJI process. Given that no simulation is required to estimate these models, it is possible to derive a closed-form solution for the option price using the characteristic function approach. An improved understanding of extreme market movements may have important implications in terms of the mispricing of derivatives.

It would also be interesting to investigate whether the ARJI model can successfully identify rational bubbles. Smith, Suchanek, and Williams (1988) have been able to create rational bubbles with a set of trading experiments. The ARJI model can be applied to these models in experimental economics to gain a better understanding of rational bubbles and market crashes.

In general, the models developed in this thesis provide a starting point for a variety of interesting and important extensions. Important contributions can be made in the areas of econometric modeling as well as forecasting tools for stock markets, foreign exchange markets, and commodity markets.

Bibliography

ADAMS, R. D., AND M. MOGHADDAM (1991): "The Risk Premia in Municipal Bond Yields: An Application of the ARCH-M model," *Journal of Macroeconomics*, 13, 725–731.

AGBEYEGBE, T. (1989): "Interest Rates and Metal Price Movements: Further Evidence," *Journal of Enviromental Economics and Managment*, 16, 184–192.

AITCHISON, J., AND C. H. HO (1989): "The Multivariate Poisson-log Normal Distribution," *Biometrica*, 76(4), 643–653.

AKGIRAY, V., AND G. G. BOOTH (1988): "Mixed Jump-Diffusion Process Modeling of Exchange Rate Movements," *Review of Economics and Statistics*, 70(4), 631 – 637.

ANDERSEN, T. G., L. BENZONI, AND J. LUND (1999): "Estimating Jump-Diffusions for Equity Returns," Kellogg Graduate School of Management, Northwestern University.

ANDERSEN, T. G., AND T. BOLLERSLEV (1998): "Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts," *International Economic Review*, 39(4), 885–905.

BABA, T., R. F. ENGLE, D. KRAFT, AND K. F. KRONER (1989): "Multivariate Simultaneous Generalized ARCH," Working Paper, University of California, San Diego.

BAILLIE, R. T., AND R. J. MYERS (1991): "Bivariate GARCH Estimation of the Optimal Commodity Futures Hedge," *Journal of Applied Econometrics*, 6, 109–124.

BAKSHI, G., C. CAO, AND Z. CHEN (1997): "Empirical Performance of Alternative Options Pricing Models," *Journal of Finance*, 52(5), 2003–2049.

BALL, C. A., AND W. N. TOROUS (1983): "A Simplified Jump Process for Common Stock Returns," *Journal of Financial and Quantitative Analysis*, 18(1), 53 – 65.

BATES, D. S. (1991): "The Crash of '87: Was it Expected? The Evidence from the Options Markets," *Journal of Finance*, 46, 1009–1044.

——— (1996): "Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options," *Review of Financial Studies*, 9(1), 69 – 107.

——— (2000): "Post-'87 Crash Fears in the S&P 500 Futures Option Market," *Journal of Econometrics*, 94, 181–238.

BAUER, R. M., F. G. NIEUWLAND, AND W. F. C. VERSCHOOR (1994): "German Stock Market Dynamics," *Empirical Economics*, 19(3), 397–418.

BEKAERT, G., AND S. F. GARY (1998): "Target Zones and Exchange Rates: Aln Empirical Investigation," *Journal of International Economics*, 45(1), 1 – 35.

BERKOWITZ, J. (2001): "Evaluating the Forecasts of Risk Models," *Journal of Business & Economic Statistics*, 19, 465–474.

BOLLERSLEV, T. (1986): "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 309–328.

BOLLERSLEV, T. (1987): "A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return," *Review of Economics and Statistics*, 69, 542–547.

BOLLERSLEV, T. (1990): "Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Model," *Review of Economics and Statistics*, pp. 498–504.

BOLLERSLEV, T., R. Y. CHOU, AND K. F. KRONER (1992): "ARCH Modelling in Finance: A Review of Theory and Empirical Evidence," *Journal of Econometrics*, 52, 5–59.

BOLLERSLEV, T., R. F. ENGLE, AND D. B. NELSON (1994): "ARCH Models," in *Handbook of Econometrics*, ed. by R. F. Engle, and D. L. McFadden, vol. 4, chap. 49, pp. 2959–3038. Elsevier Science.

BOLLERSLEV, T., R. F. ENGLE, AND J. M. WOOLDRIDGE (1988): "A Capital Asset Pricing Model with Time-Varying Covariances," *Journal of Political Economy*, 96(1), 116–131.

BOLLERSLEV, T., AND J. WOOLDRIDGE (1992): "Quasi-Maximum Likelihood Estimation and Inference in Dynamics Models with Time-Varying Covariances," *Econometric Review*, 11, 143–172.

BOX, G. E., AND G. C. TIAO (1973): *Bayesian Inference in Statistical Analysis.* Addison-Wesley, Cambridge Massachusetts.

BREEDEN, D., M. R. GIBBONS, AND R. LITZENBERGER (1989): "Empirical Tests of Consumption Oriented CAPM," *Journal of Finance*, 44, 231–262.

BROCK, W., J. LAKONISHOK, AND B. LEBARON (1992): "Simple Technical Trading Rules and the Stochastic Properties of Stock Returns," *Journal of Finance*, 47(5), 1731 – 1764.

BURT, J., F. R. KAEN, AND G. G. BOOTH (1977): "Foreign Market Efficiency Under Flexible Exchange Rates," *Journal of Finance*, 32, 1325–1330.

CAMPBELL, J. T. (1934): "The Poisson Correlation Function," *Proceedings of the Edinburgh Mathematical Society, Series 2*, 4, 18–26.

CHERNOV, M., A. R. GALLANT, E. GHYSELS, AND G. TAUCHEN (1999): "A New Class of Stochastic Volatility Models with Jumps: Theory and Estimation," Working Paper, CIRANO, the University of Montreal.

CLARK, P. K. (1973): "A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices," *Econometrica*, 41(1), 135 – 155.

CRAINE, R., A. LOCHSTOER, AND K. SYRTVEIT (2000): "Estimation of a Stochastic-Volatility Jump Diffusion Model," University of California at Berkeley.

DAOROGNA, M., R. GENÇAY, U. MÜLLER, R. OLSEN, AND O. PICTET (2001): *An Introduction to High Frequency Finance*. Academic Press.

DAS, S. R. (1998): "Poisson-Gaussian Processes and the Bond Market," NBER Working Paper 6631.

DAS, S. R., AND R. K. SUNDARAM (1997): "Taming the Skew: Higher-Order Moments in Modeling Asset Price Processes in Finance," NBER Working Paper 5976.

——— (1999): "Of Smiles and Smirks: A Term Structure Perspective," *Journal of Financial and Quantitative Analysis*, 34(2), 211–240.

DAVIES, R. B. (1987): "Hypothesis Testing when a Nuisance Parameter is Present Only Under the Alternative," *Biometrika*, 74, 33–43.

DIEBOLD, F. X., T. A. GUNTHER, AND A. S. TAY (1998): "Evaluating Density Forecasts with Application to Financial Risk Management," *International Economic Review*, 39(4), 863–883.

DIEBOLD, F. X., AND M. NERLOVE (1989): "The Dynamics of Exchange Rate Volatility: A Multivariate Latent Factor ARCH Model," *Journal of Applied Econometrics*, 4, 1–21.

DOMOWITZ, I., AND C. S. HAKKIO (1985): "Conditional Variance and the Risk Premium in Foreign Exchange Market," *Journal of International Economics*, 19, 47–65.

DORNBUSCH, R. (1976): "Expectations and Exchange Rate Dynamics," *Journal of Political Economy*, 84, 1161–1176.

DROST, F. C., T. E. NIJMAN, AND B. J. M. WERKER (1996): "Estimation and Testing in Models Containing Both Jumps and Conditional Heteroskedasticity," Working Paper, Tilburg University.

EDWARDS, C. B., AND J. GURLAND (1961): "A Class of Distributions Applicable to Accidents," *Journal of American Statistical Association*, pp. 503–517.

ENGLE, R. F. (1982): "Autoregressive Conditional Heteroskedasticity with Estimates of the UK inflation," *Econometrica*, 50, 987–1008.

ENGLE, R. F., AND T. P. BOLLERSLEV (1986): "Modelling the Persistence of Conditional Variances," *Econometric Review*, 5, 1–50.

ENGLE, R. F., V. K. NG, AND M. ROTHSCHILD (1990): "Asset Pricing with a Factor-ARCH Covariance Structure: Empirical Estimates for Treasury Bills," *Journal of Econometrics*, 45, 213–237.

ERAKER, B., M. S. JOHANNES, AND N. G. POLSON (1999): "Return Dynamics in Continuous-Time with Jumps to Volatility and Returns," Graduate School of Business, University of Chicago.

EVANS, M. D. D., AND R. K. LYONS (2001): "Order Flow and Exchange rate Dynamics," Haas School of Business, UC Berkeley, Working Paper.

FAMA, E. F. (1965): "The Behavior of Stock Market Prices," *Journal of Business*, 38, 34–105.

——— (1970): "Efficient Capital Markets: A Review of Theory and Empirical Work," *Journal of Finance*, 30, 383–417.

FISCHER, A. (1988): "Money Announcements and the Risk Premium," *Economics Letters*, 27, 155–158.

FORTUNE, P. (1999): "Are Stock Returns Different Over Weekends? A Jump Diffusion Analysis of the Weekend Effect," *New England Economic Review*, pp. 3 – 19.

FRENKEL, J. A. (1976): "A Monetray Approach to the Exchange Rate: Doctrinal Aspects and Empirical Evidence," *Scandinavian Journal of Economics*, 78, 200–224.

——— (1979): "On the Mark: A Theory of Floating Exchange Rates Based on Real Interest Differentials," *American Economic Review*, 69, 610–622.

FRIEDMAN, D., AND S. VANDERSTEEL (1982): "Short-Run Fluctuations in Foreign Exchange Rates," *Journal of International Economics*, 13, 171–186.

GALLANT, A. R., D. HSIEH, AND G. TAUCHEN (1997): "Estimation of Stochastic Volatility Models with Diagnostics," *Journal of Econometrics*, 81, 159–192.

GALLANT, A. R., AND G. TAUCHEN (1989): "Semi-Non-Parametric Estimation of Conditionally Constrained Heterogeneous Process: Asset Pricing Applications," *Econometrica*, 57, 1091–1120.

GAUDET, G., AND P. HOWITT (1989): "A Note on Uncertainty and Hotelling Rule," *Journal of Environmental Economics and Managment*, 16, 80–86.

GAUDET, G., AND A. M. KHADR (1991): "The Evolution of Natural Resource Prices Under Stochastic Investment Opportunities: An Intertemporal Asset Pricing Approach," *International Economic Review*, 32, 441–445.

GENÇAY, R., G. BALLOCCHI, M. DAOROGNA, R. OLSEN, AND O. PICTET (2002): "Real-Time Trading Models and the Statistical Properties of Foreign Exchange Rates," *forthcoming in International Economic Review.*

GENÇAY, R., M. DAOROGNA, R. OLSEN, AND O. PICTET (2002): "Real-Time Foreign Trading Models and Market Behavior," *forthcoming in Journal of Economic Dynamics and Control.*

GOURIEROUX, C., A. MONFORT, AND A. TROGNON (1984): "Pseudo Maximum Likelihood Methods: Applications to Poisson Models," *Econometrica*, 52(3), 701–720.

HAMDAN, M. A., AND H. A. AL-BAYYATI (1969): "A Note on the Bivariate Poisson Distribution," *The American Statistician*, 23(4), 32–33.

HANSEN, B. E. (1996): "Inference When a Nuisance Parameter is Not Identified Under the Null Hypothesis," *Econometrica*, 64(2), 413–430.

HANSEN, L., AND R. HODRICK (1983): "Risk Averse Speculation in the Forward Foreign Exchange Market: An Econometric Analysis of Linear Models," in Frenkel, J.A. eds. Exchange Rates and International Macroeconomics.

HANSEN, L., AND S. RICHARD (1987): "The Role of Conditioning Information in Deducing Testable Restrictions implied by Dynamic Asset Pricing Models," *Econometrica*, 55, 587–614.

HEAL, G., AND M. BARROW (1980): "The Relationship Between Interest Rates and Metal Price Movements," *Review of Economics Studies*, 47, 161–181.

HIGGINS, M. L., AND A. K. BERA (1992): "A Class of Nonlinear ARCH Models," *International Economic Review*, 33, 137–158.

HOLGATE, P. (1964): "Estimation for the Bivariate Poisson Distribution," *Biometrica*, 51, 241–245.

HOTELLING, H. (1931): "The Economics of Exhausible Resources," *Journal of Political Economy*, 39, 137–175.

HSIEH, D. (1989a): "Modeling Heteroscedasticity in Daily Foreign Exchange Rates," *Journal of Business & Economic Statistics*, 7(3), 307–317.

——— (1989b): "Testing for Nonlinearity in Daily Foreign Exchange Rate Changes," *Journal of Business*, 62, 339–368.

JORION, P. (1988): "On Jump Processes in the Foreign Exchange and Stock Markets," *Review of Financial Studies*, 1(4), 427–445.

JUNG, R. C., AND R. WINKELMANN (1993): "Two Aspects of Labour Mobility: A Bivariate Poisson Regression Approach," *Empirical Economics*, 18, 543–556.

KENDALL, M., AND A. STUART (1969): *The Advanced Theory of Statistics*, vol. 1. Charles Griffin, London.

KHALAF, L., J.-D. SAPHORES, AND J.-F. BILODEAU (2000): "Simulation-Based Exact Tests with Unidentified Nuisance Parameters Under the Null Hypothesis: the Case of Jumps Tests in Models with Conditional Heteroskedasticity," Laval University, Working Paper 2000.

KRAUTKRAEMER, A. J. (1998): "Nonrenewable Resource Scarcity," *Journal of Economic Literature*, 36, 2065–2107.

LEE, T. K. (1988): "Does Conditional Covariance or Conditional Variance Explain Time Varying Risk Premia in Foreign Exchange Returns?," *Economics Letters*, 27, 371–373.

LEE, T. K. Y., AND Y. K. TSE (1991): "Term Structure of Interest Rate in Singapore Asian Dollar Market," *Journal of Applied Econometrics*, 6, 143–152.

LIN, S., J. KNIGHT, AND S. SATCHELL (1999): "Modelling Intra-day Equity Prices and Volatility Using Information Arrivals - A comparative Study of Different Choices of Information Proxies," *in Financial Market Tick by Tick.*

LJUNG, G., AND E. BOX (1978): "On a Measure of Lack of Fit in Time Series Models," *Biometrica*, 65, 297–303.

MANDELBROT, B. (1963): "The Variation of Certain Speculative Prices," *Journal of Business*, 36, 394–419.

MARSHALL, A. W., AND I. OLKIN (1985): "A Family of Bivariate Distribution Generated by the Bivariate Bernoulli Distribbution," *Journal of the American Statistical Association*, 80(390), 332–338.

MARTIZ, J. S. (1952): "Note on a Certain Family of Discrete Distribution," *Biometrica*, 39, 196–198.

MAYER, W. J., AND W. F. CHAPPELL (1992): "Determinants of Entry and Exit: An Application of the Compounded Bivariate Poisson Distribution to U.S. Industries," *Southern Economic Journal*, 58(3), 770–778.

MCCURDY, T. H., AND I. G. MORGAN (1988): "Testing the Martingale Hypothesis in the Deutsche Mark Futures with Models Specifying the Form of Heteroskedasticity," *Journal of Applied Econometrics*, 3, 187–203.

——— (1991): "Test for a Systematic Risk Component in Deviations from Uncovered Interest Rate Parity," *Review of Economic Studies*, 58, 587–602.

——— (1992): "Evidence of Risk Premiums in Foreign Currency Futures Markets," *Review of Financial Studies*, 5(1), 65–83.

——— (1993): "Intertemporal Risk in Foreign Currency Markets," in *The Exchange Rate and the Economy*, pp. 325–355. Bank of Canada.

MCDONALD, J. B., AND W. K. NEWEY (1988): "Partially Adaptive Estimation of Regression Models vis the Generalized t Distribution," *Econometric Theory*, 4, 428–457.

MEESE, R. A., AND K. ROGOFF (1983): "Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample," *Journal of International Economics*, 14, 3–24.

MERTON, R. C. (1980): "On Estimating the Expected Return on the Market: An Exploratory Investigation," *Journal of Financial Economics*, 8, 323–361.

M'KENDRICK, A. G. (1926): "Applications of Mathematics to Medical Problems," *Proceedings of the Edinburgh Mathematical Society*, 44, 98–130.

NEELY, C. J. (1999): "Target Zones and Conditional Volatilty: The Role of Realignments," *Journal of Empirical Finance*, 6(2), 177 – 192.

NELSON, D. (1988): "Conditional Heteroskedasticity in Asset Returns: A New Approach," PhD dissertation, M.I.T., Department of Economics.

NIEUWLAND, F., W. VERSHCHOOR, AND C. WOLFF (1994): "Stochastic Trends and Jumps in EMS Exchange Rates," *Journal of International Money and Finance*, 13, 699–727.

PALM, F. C., AND P. J. G. VLAAR (1997): "Simple Diagnostic Procedures for Modeling Financial Time Series," *Allgemeines Statistisches Archiv*, 81(1), 85–101.

PARKINSON, M. (1980): "The Extreme Value Method for Estimating the Variance of the Rate of Return," *Journal of Business*, 53(1), 61–65.

PRESS, S. J. (1967): "A Compound Events Model for Security Prices," *Journal of Business*, 40, 317 – 335.

ROGALSKI, R. J., AND J. D. VINSO (1978): "Empirical Properties of Foreign Exchange Rates," *Journal of International Business Studies*, 9, 69–79.

ROSENBLATT, M. (1952): "Remarks on a Multivariate Transformation," *Annals of Mathematical Statistics*, 23, 470–472.

SAPHORES, J.-D., L. KHALAF, AND D. PELLETIER (2000): "On Jumps and ARCH Effects in Natural Resource Prices: An Application to Stumpage Prices from Pacific Northwest National Forests," Laval University, Working Paper 2000.

SLADE, M. (1982): "Trends in Natural Resource Commodity Prices: An Analysis of the Time Domain," *Journal of Environmental Economics and Management*, 9, 122–137.

——— (1988): *Pricing of Metals.* Kingston: Centre for Resource Studies.

——— (1991): "Market Structure, Marketing Method, and Price Instability," *Quarterly Journal of Economics*, 106, 1309–1340.

SLADE, M., AND H. THILLE (1995): "Hotelling Confronts CAPM: A Test of the Theory of Exhaustable Rsources," *Canadian Journal of Economics*, 30, 685–708.

SMITH, V. (1981): "The Empirical Relevance of Hotelling's Model for Natural Resource," *Resources and Energy*, 3, 105–117.

SMITH, V. L., G. L. SUCHANEK, AND A. W. WILLIAMS (1988): "Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets," *Econometrica*, 56(5), 1119–1151.

SO, M. K. P., K. LAM, AND W. K. LI (1999): "Forecasting Exchange Rate Volatility Using Autoregressive Random Variance Model," *Applied Financial Economics*, 9, 583–591.

SUSMEL, R. (1994): "Hourly Volatility Spillover Between International Equity Markets," *Journal of International Money and Finance*, 13, 3–25.

TEICHER, H. (1954): "On the Multivariate Poisson Distribution," *Skandinavisk Aktuarietidskrift*, 37, 1–9.

TUCKER, A. L., AND L. POND (1988): "The Probability Distribution of Foreign Exchanges: Tests of Candidate Processes," *Review of Economics and Statistics*, 70(4), 638 – 647.

VLAAR, P. J. G., AND F. C. PALM (1993): "The Message in Weekly Exchange Rates in the European Monetary System: Mean Reversion, Conditional Heteroscedasticity, and Jumps," *Journal of Business & Economic Statistics*, 11(3), 351–360.

WEST, K. D., AND D. CHO (1995): "The Predictive Ability of Several Models of Exchange Rate Volatility," *Journal of Econometrics*, 69, 367–391.

WESTERFIELD, J. M. (1977): "An Examination of Foreign Exchange Risk Under Fixed and Floating Rate Regimes," *Journal of International Economics*, 7, 181–200.

YOUNG, D., AND D. L. RYAN (1996): "Empirical Testing of a Risk-Adjusted Hotelling Model," *Resource and Energy Economics*, 18, 265–289.