Risks in Commodity and Currency Markets

Miloš Božović

TESI DOCTORAL UPF / 2009

DIRECTOR DE LA TESI

Prof. Branko Urošević

TUTOR DE LA TESI

Prof. Xavier Freixas

UNIVERSITAT POMPEU FABRA

Dipòsit legal:

ISBN:

to the memory of my father

Acknowledgment

This thesis was written during my graduate studies at the Department of Economics and Business, University Pompeu Fabra, in the years 2005–2009. I am grateful to the Department for providing a stimulating research environment and the possibilities for attending excellent courses, seminars, workshops and conferences.

I owe a substantial debt to my advisor Branko Urošević, whose inspiring influence got me interested in the field of financial economics. His guidance, advices and continuous encouragements were an invaluable source of enthusiasm for all these years. I also owe much gratitude to Xavier Freixas for his support throughout my journey towards the Ph.D.

I am grateful to Manuel Moreno, whose cogent suggestions truly improved my thesis and helped its early-stage progress. In addition, the remarkable courses given by him and Jon Danielsson taught me many of the tools I later used in my research.

I have received valuable comments and suggestions from Francesco Corielli, Andras Fulop, Eric Ghysels, Benjamin Golež, Carlos Gonzales, José Marín, Lionel Martellini, Loriana Pelizzon, Francisco Peñaranda, Gonzalo Rubio, Alessandro Sbuelz, Jaume Ventura and Tanja Veža.

I also want to thank Marta Araque, Marta Aragay, Gemma Burballa and Raquel Iglesias for all their help in the organizational matters.

I should not forget my former students, who kept me attentive throughout these years and taught me that the most profound questions are often the simplest ones.

Special thanks goes to Rhiannon, Juan Manuel and Blaž, with whom I had shared the office for several years, enjoying lots of inspiring moments and countless academic and not-so-academic discussions. Thanks to Raša for many inspiring ideas and to Ognjen for sharing various thoughts during the coffee brakes. It is them and many other fellow Ph.D. students who made this endeavor much more enjoyable.

The selfless support of my sister Jelena made this thesis possible. I am indebted to her for the love, understanding and willingness to bear the burden of my absence.

I am eternally grateful to my late parents for being the best role models one could possibly have.

Finally, thanks to you Jasmina, for putting up with me and my distractions for all these years. Your love and inspiration were truly indispensable.

Miloš Božović, Barcelona, March 2009

Abstract

This thesis analyzes market risk factors in commodity and currency markets. It focuses on the impact of extreme events on the prices of financial products traded in these markets, and on the overall market risk faced by the investors. The first chapter develops a simple two-factor jump-diffusion model for valuation of contingent claims on commodities in order to investigate the pricing implications of shocks that are exogenous to this market. The second chapter analyzes the nature and pricing implications of the abrupt changes in exchange rates, as well as the ability of these changes to explain the shapes of option-implied volatility smiles. Finally, the third chapter employs the notion that key results of the univariate extreme value theory can be applied separately to the principal components of ARMA-GARCH residuals of a multivariate return series. The proposed approach yields more precise Value at Risk forecasts than conventional multivariate methods, while maintaining the same efficiency.

Resumen

El objetivo de esta tesis es analizar los factores del riesgo del mercado de las materias primas y las divisas. Está centrada en el impacto de los eventos extremos tanto en los precios de los productos financieros como en el riesgo total de mercado al cual se enfrentan los inversores. En el primer captulo se introduce un modelo simple de difusión y saltos (*jump-diffusion*) con dos factores para la valuación de activos contingentes sobre las materias primas, con el objetivo de investigar las implicaciones de shocks en los precios que son exógenos a este mercado. En el segundo capítulo se analiza la naturaleza e implicaciones para la valuación de los saltos en los tipos de cambio, así como la capacidad de éstos para explicar las formas de sonrisa en la volatilidad implicada. Por último, en el tercer capítulo se utiliza la idea de que los resultados principales de la Teoria de Valores Extremos univariada se pueden aplicar por separado a los componentes principales de los residuos de un modelo ARMA-GARCH de series multivariadas de retorno. El enfoque propuesto produce pronósticos de Value at Risk más precisos que los convencionales métodos multivariados, manteniendo la misma eficiencia.

Contents

Foreword

The motivation of this thesis is to explore the market risk factors in commodity and currency markets. In particular, it focuses on the impact of extreme events on the prices of financial products traded in these markets. It is also an attempt to develop a comprehensive framework for the assessment of the overall market risk faced by investors that would include the risk of extreme losses.

The thesis consists of three essays. In Chapter 1, I develop a simple two-factor jump-diffusion model for pricing of contingent claims on commodities in order to investigate the pricing implications of shocks that are exogenous to the market. The model is constructed such that it explicitly accounts for the key features of commodity prices: the mean reversion, the correlation with the risk-free interest rate and, most importantly, the possibility of abrupt jumps that affect the prices substantially. For the proposed model, I provide a closed form solution for the price of a forward and a futures contract on a commodity, as well as a semi-closed form solution for the price of European options on commodity futures. Furthermore, I derive an expression for the optimal hedging ratio of a dynamic futures hedge. The model parameters are calibrated by adapting the maximum likelihood technique to account for jumps in the spot price. The empirical results based on the futures prices and treasury yields indicate that, in addition to the principal risk factor related to everyday supply and demand patterns, which is commodity spot price diffusion, other factors cannot be neglected. Specifically, the market data for some commodities implies that investors require substantial premia for the exposure to the jump risk. These additional premia are comparable to the one originating from the diffusion and higher than the interest rate risk premia.

Chapter 2 is related to the role of jump risk in foreign exchange rates. It analyzes the nature and pricing implications of jumps, the abrupt changes in the exchange rates. I propose a general stochastic-volatility jump-diffusion type of model of exchange rate dynamics that contains several popular models as its special cases. I use the efficient method of moments to estimate the model parameters from the spot exchange rates of Euro, British Pound, Japanese Yen and Swiss Franc with respect to the U.S. Dollar. The results indicate that any reasonably descriptive continuous-time model of exchange rates must allow for jumps with a bimodal distribution of jump sizes and in some cases jump frequencies that depend on volatility. In the second part of the chapter, I investigate the option pricing implications of jumps. Although the ex-post estimates of jump probabilities show that jumps occur irregularly and rarely, the jump component is crucial for explaining the shapes of implied volatility "smiles". The risk premia calculated from the cross-sectional currency options data suggest that the exchange-rate jump risk appears to be priced by the market.

The risk of extreme events in currency markets can be also assessed using an alternative approach. In Chapter 3, I develop an efficient method based on multivariate extreme value theory to measure the market risk of a portfolio. The approach employs the notion that some key results of the univariate extreme value theory can be applied separately to a set of orthogonal random variables, provided they are independent and identically distributed. Such random variables can be constructed from the principal components of ARMA-GARCH conditional residuals of a multivariate return series. The model's forecasting ability is then tested on a portfolio of foreign currencies. The results indicate that the generalized Pareto distribution of peeks over treshold of residuals performs well in capturing extreme events. In particular, model backtesting shows that the proposed multivariate approach yields more precise Value at Risk forecasts than the usual methods based on conditional normality, conditional t-distribution or historical simulation, while maintaining the efficiency of conventional multivariate methods.

Chapter 1

Valuation of Contingent Claims on Commodities under Mean-Reverting Jump Diffusion

1.1 Introduction

The stochastic behavior of commodity prices plays a central role in valuation of financial contingent claims on commodities and investments to extract them. As pointed out by Schwartz (1997), models that assume constant interest rates and constant convenience yields fail to capture one of the basic properties in the behavior of commodity prices – the mean reversion. The mean-reverting nature of commodity prices is now a well-established empirical fact.[1] It also has a firm microeconomic ground. Namely, in an equilibrium setting one would expect that when prices are relatively high, supply will increase since the suppliers with higher input costs will enter the market, thereby putting a downward pressure on prices. Conversely, when prices are relatively low, supply will decrease since some of the suppliers with higher input costs will exit the market, thereby putting an upward pressure on prices.

[1]See, for example, Pindyk and Rubinfeld (1991) or Pilipovic (1998).

This specific interaction of relative prices and supply induces the mean reversion. A series of articles discusses the mean-reverting property of commodity prices. See, for example, Gibson and Schwartz (1990), Brennan (1991), Cortazar and Schwartz (1994), Bessembinder, Coughenour, Seguin, and Smoller (1995), Schwartz (1997), Pindyk (2001), Cortazar and Schwartz (2003), and Lei and Fox (2004).

A basic model for commodity pricing takes into account only one risk factor – the spot price. Since geometric Brownian motion (GBM) has unboundedly increasing variance as time horizon increases, it cannot capture the mean-reverting property. Schwartz (1997) (with a reference to unpublished manuscript by Ross (1995)) proposes a process of the Ornstein-Uhlenbeck type for the logarithm of commodity spot price. Earlier, Gibson and Schwartz (1990) suggested a two-factor model for pricing contingent claims on oil. In their model, the first stochastic factor is the spot price of oil, described by the GBM. The second stochastic factor is the convenience yield, which can be defined as the benefit obtained from holding the spot (i.e. physical) commodity that is not obtained from holding the futures contract (see Brennan (1991)). The benefits typically include the ability to keep production process running or to profit from temporary local shortages of the commodity. Although such benefits depend on the individual storing the commodity, the equilibrium futures price will depend on an equilibrium convenience yield obtained from competition between potential storers. Apart from benefits and storage costs, the convenience yield also implicitly accounts for the cost of insuring the commodity. In the Gibson and Schwartz (1990) model, the convenience yield is considered to be mean reverting and positively correlated with the spot price. The extension of this model allows for a stochastic interest rate (see Schwartz (1997)). As Crosby (2005) points out, these two- and three-factor models, although based on a GBM for the spot price, implicitly account for the mean reversion property as long as both interest rate and convenience yield are governed by a mean reverting process. The best feature of these models is their full tractability, since they allow for closed form solutions for futures prices and for a linear relation between the logarithm of futures prices and the underlying factors.

However, simplicity of factor models is not without a cost: they fail to capture another very important feature. Namely, the price processes for many commodities are influenced by the arrival of important new information that has more than a marginal effect on price. Such abrupt changes were first considered by Merton (1976) in a model for stock prices. By including a Poisson jump component to the usual GBM, Merton obtained a closed form solution for the price of a European call option on a dividend-paying stock. Hilliard and Reis (1998) apply the same model for the spot prices of commodities within the framework of a three-factor model with stochastic convenience yields and interest rates. Jump-diffusion processes have also been used in models for commodity prices in Deng (1998), Clewlow and Strickland (2000), and Benth, Ekeland, Hauge, and Nielsen (2003). A potential drawback of all these models is that they explicitly specify the dynamics of convenience yield. As convenience yield is not directly observable, there is no empirical evidence that would support any of the chosen dynamics. Therefore, any results derived from such models may rely too much on correct specification for the convenience yield. There is also no evidence that different classes of commodities would follow the same convenience yield dynamics.

Crosby (2005) proposed a general model for pricing of commodity derivatives that incorporates wide range of empirical facts and avoids explicit modeling of the convenience yield. The model specifies stochastic processes for the futures price and the interest rate, accounting for features such as jumps with a magnitude that depends on time to delivery, volatility skews and seasonality, or time-varying long-run equilibrium level in the term structure. Although very intuitive, this model cannot deliver any closed-form solution for values of derivative contracts. Its application is therefore restricted to numerical procedures such as Monte Carlo simulation. Also, no calibrating techniques were developed for it to date.

In this chapter I develop a simple two-factor model for pricing contingent claims on commodities. I assume a frictionless arbitrage-free market that trades continuously. I propose a joint stochastic process for the commodity spot price and the short interest rate, taking into account possible correlation between them. The spot price is considered to follow a process of a geometric Ornstein-Uhlenbeck type, with additional compensated Poisson jumps. On the other hand, the term structure of interest rates is modeled by a Vasicek process. In this way, I am able to account for the key features of the contingent claim pricing factors explicitly. These key factors are: (1) mean reversion in commodity spot prices; (2) abrupt jumps with more than a marginal effect on commodity prices; (3) mean reversion in the interest rate; (4) correlation between marginal changes in the interest rate and the commodity spot prices. Seasonality, which is another prominent feature of commodity prices, is not explicitly modeled. Rather, I calibrate the model using deseasonalized time series.

The model allows for a closed form solution for the price of a forward and a future contract, as well as a series expansion for the price of a European option on futures. Furthermore, I derive an expression for the optimal hedging ratio of a dynamic futures hedge involving calendar spreads. Under proposed setting, the returns on commodity prices will be affected by three distinct risk factors: the diffusive shocks in the spot prices, the diffusive shocks in the interest rate, and the price jumps.

Even though I do not use the convenience yield as an explicit pricing factor, it is implicitly embedded in the model. Casassus and Collin-Dufresne (2005) use a similar approach, although in a different framework. In this way, the model can be calibrated more efficiently, by using only two time series related to assets that are actually traded – the prices of futures contracts and the yields on Treasury bills. Usually, factor models of this sort are calibrated with the Kalman filter technique. By taking the full advantage of the analytical tractability of the model developed in this chapter, I introduce much simpler estimation procedure that is based on the maximum likelihood technique and does not require filtering. I use historical data for yields on U.S. 3-Month Treasury Bills and prices of futures contracts on several commodities during past 16 years to calibrate the model. The commodity data consists of weekly observations for six exchange-traded commodities: Brent crude oil, natural gas, copper, gold, wheat and pork bellies. The empirical results of this chapter indicate that all three risk factors are significant for commodity pricing. They are also priced differently. The jump risk (which occurs due to combined uncertainty in timing and magnitude of the jumps) carries an important part of the overall risk premium.

The remainder of the chapter is organized as follows. The valuation model is developed in Section 1.2. Section 1.3 provides closed form solutions for the values of various contingent claims (forward and futures contracts, implied convenience yield, and European option on futures) and calculates the optimal hedging ratio. Section 1.4 describes the procedure for empirical estimation of the model parameters and reports the obtained values, along with associated market prices of risk. Section 1.5 discusses model implications and extensions, and concludes.

1.2 Valuation Model

I construct a two-factor model for contingent-claim valuation. The first factor is the spot price of commodity, the second is the term structure of interest rates. The spot price of commodity is assumed to follow a process of geometric Ornstein-Uhlenbeck (GOU) type, with discrete jumps. Following Clewlow, Strickland, and Kaminski (2001), I add a Poisson jump component to the Schwartz (1997) model and assume that the spot price dynamics under physical (i.e., data-generating) probability measure $\mathbb{P}$ is governed by the following stochastic differential equation:

$$\frac{\mathrm{d}S_t}{S_t} = \left[a\left(m - \ln S_t\right) - \lambda k\right]\mathrm{d}t + \sigma \mathrm{d}W_t + (U_t - 1)\mathrm{d}q_t. \tag{1.1}$$

Here, $\mathrm{d}W$ is the standard Wiener process, while $\mathrm{d}q$ is the Poisson process with parameter λ. Using the same specification as in Merton (1976), I suppose the following characterization of the Poisson process: The jumps are serially uncorrelated; the probability of a jump occurring once during a time interval of length $\mathrm{d}t$ is $\lambda \mathrm{d}t + \mathrm{O}(\mathrm{d}t)$; the probability of a jump not occurring during the same time interval is $1 - \lambda \mathrm{d}t + \mathrm{O}(\mathrm{d}t)$; the probability of a jump occurring more than once in $\mathrm{d}t$ is $\mathrm{O}(\mathrm{d}t)$. Conditionally on the Poisson event occurring, a new random variable U_t measuring the size of its impact is drawn from a known distribution. The size of this random jump is equal to $U_t - 1$, with expected value $\varepsilon_t (U_{t+} - 1) = k$. Here, ε denotes the expectation with respect to the distribution of U. The contribution of each jump to the price change between instants t and $t + \mathrm{d}t$ is thus $S_t U_t$. Subtraction of $\lambda k \mathrm{d}t$ in equation (1.1) centers $\mathrm{d}q$ around its expectation, and hence the expected return from t to $t + \mathrm{d}t$ is simply equal to $a (m - \ln S_t)\, \mathrm{d}t$, so the process is clearly mean reverting. The parameter $a > 0$ measures the speed of mean reversion to the long run mean log price, m.

The total change in the spot price is therefore composed of two types of changes. The first one is affected by temporary imbalances between supply and demand or any new information that causes a marginal change in the price level. This component is modeled by a mean reverting process of GOU type with a standard Brownian diffusion. The second type of change comes from arrival of important new information that has more than a marginal effect on price. This change is typically characterized by low frequency and sudden occurrence. Typical examples are shocks due to wars or natural disasters, supply shocks in markets with cartelized commodities (such as oil), but also endogenous commodity market shock that cause abnormal returns in a very short period of time.[2] I will assume that, due to a different nature of the mechanism behind, the processes $\mathrm{d}W$ and $\mathrm{d}q$ are uncorrelated.

[2]The episode of Amaranth's $6 billion loss is perhaps the best example of the latter. Much of this loss came from positions the fund had in natural gas futures, which plummeted around September 15, 2006.

It is worth noting that jumps in the spot price lead to distributions that have degrees of skewness and kurtosis different from those of the log-normal distribution. These distributions can lead to values of derivative contracts that differ considerably from those obtained with the corresponding GBM process.

The second factor I use for pricing is the short interest rate. We conjecture that the term structure follows the Vasicek (1977) model,

$$\mathrm{d}r_t = b\left(\bar{r} - r_t\right)\mathrm{d}t + \theta \mathrm{d}Z_t, \tag{1.2}$$

where $\mathrm{d}Z$ is a standard Wiener process different from the one in equation (1.1). However, I will assume that the two diffusions are correlated in general. I take this into account by setting $\mathrm{d}W_t\mathrm{d}Z_t = \rho\mathrm{d}t$. The process in equation (1.2) is assumed to be AOU merely for the sake of simplicity, since in this way we are able to cast our model easily into a vector form (see Section 4). Of course, there is a multitude of plausible models for the short interest rate. For example, Cox, Ingersoll, and Ross model (see Cox, Ingersoll, and Ross (1985)), or a model from the Heath, Jarrow, and Morton family (see Heath, Jarrow, and Morton (1992)), have been widely used throughout the literature. Since I tend to focus here on commodity price dynamics, it is reasonable to conjecture that the choice of interest rate model is not so essential.

Under assumed dynamics, given by equations (1.1) and (1.2), any position in commodity derivative contract will be affected by four sources of risk. These are the spot price diffusive risk associated with the diffusion process $\mathrm{d}W$, interest rate diffusive risk driven by $\mathrm{d}Z$, spot price jump-time risk whose origin is the random timing of jump events $\mathrm{d}q$, and finally spot price jump-size risk driven by U.

The usual no-arbitrage approach to valuation of contingent claims involves calculation of expected values in a risk-neutral world.[3] According to the Fundamental Theorem of Asset Pricing, the market will be arbitrage-free if and

[3] An alternative approach is preference-based equilibrium pricing.

only if there exists an equivalent martingale measure (see Harrison and Pliska (1981) or Elliot and Kopp (2005)). Furthermore, if the market is complete this measure will be unique. In general, the two additional sources of uncertainty, namely the random spot price jump times and random jump sizes, make the market incomplete with respect to the risk-free bank account, the commodity spot contract, and the finite number of option contracts. Consequently, the state-price density (the "pricing kernel") will not be unique.[4] However, by modeling the jump process as in equation (1.1) we entangle two jump-risk components into one, thereby restoring the uniqueness of the risk-neutral measure. To express the price dynamics under such a measure $\mathbb{P}^*$, the stochastic processes in equations (1.1) and (1.2) have to be transformed, respectively, to

$$\frac{\mathrm{d}S_t^*}{S_t} = \left[a\left(m^* - \ln S_t\right) - \lambda^* k^*\right]\mathrm{d}t + \sigma \mathrm{d}W_t^* + (U_t - 1)\mathrm{d}q_t^*, \tag{1.3}$$

and

$$\mathrm{d}r_t^* = b\left(\bar{r}^* - r_t\right)\mathrm{d}t + \theta \mathrm{d}Z_t^*. \tag{1.4}$$

Expressions (1.3) and (1.4) are derived in Appendix A. The transformed Wiener processes satisfy $\mathrm{d}W_t^*\mathrm{d}Z_t^* = \rho \mathrm{d}t$. The parameters $\bar{x}^*$, $\bar{r}^*$, λ^*, and k^* are related to $\bar{x}$, $\bar{r}$, λ, and k via

$$m^* = m - \frac{\sigma}{a}\left(\xi_1 + \rho\xi_2\right) \tag{1.5}$$

$$\bar{r}^* = \bar{r} - \frac{\theta}{b}\left(\rho\xi_1 + \xi_2\right) \tag{1.6}$$

$$\lambda^* = \lambda e^{\xi_3} \tag{1.7}$$

$$1 + k^* = (1 + k)e^{\xi_4 \omega^2} \tag{1.8}$$

where ξ_1 though ξ_4 are parameters that measure the market prices of risk and ω^2 is the variance of $\ln U$ (cf. Section 3 and Appendix A). By comparing the risk-neutral processes (1.3) and (1.4) to their $\mathbb{P}$-equivalents, (1.1) and (1.2), we can see that the pricing factors appreciate with the following risk premia:

[4] In a preference-based model, the state-price density arises from marginal rates of substitution evaluated at equilibrium consumption streams.

- Premium for the spot price diffusive risk: $\sigma(\xi_1 + \rho\xi_2)/a$.
- Premium for the interest rate diffusive risk: $\theta(\rho\xi_1 + \xi_2)/b$.
- Overall premium for the jump risk: $\lambda k - \lambda^* k^*$.

Finally, by applying Itô's lemma to $x_t = \ln S_t$ we can rewrite equation (1.1) in the form of a Vasicek process with a jump component. The dynamics of this component under the risk neutral measure $\mathbb{P}^*$ is given by

$$\mathrm{d}x_t^* = \left[a\left(\bar{x}^* - x_t\right) - \lambda^* k^*\right]\mathrm{d}t + \sigma \mathrm{d}W_t^* + \ln U_t \mathrm{d}q_t^*, \tag{1.9}$$

where $\bar{x}^* = m^* - \sigma^2/2a$.

1.3 Prices of Contingent Claims

1.3.1 Prices of Forward and Futures Contracts

In this section I derive analytical results for prices of basic contingent claims on commodities in the framework of the proposed two-factor model. I start by computing the price at t of a unit (default-free) discount bond maturing at T:

$$B(t,T) = \mathbb{E}_t^*\left[\exp\left(-\int_t^T r_s \mathrm{d}s\right)\right]. \tag{1.10}$$

The integral

$$\int_t^T r_s \mathrm{d}s = (T-t)\bar{r}^* + \frac{1 - e^{-b(T-t)}}{b}(r_t - \bar{r}^*) + \theta \int_t^T \frac{1 - e^{-b(T-s)}}{b}\mathrm{d}Z_s^* \tag{1.11}$$

is a normal random variable conditionally on information available at t, being the sum of a constant, a measurable random variable, and an Itô integral. Therefore, the expression inside the expectation in equation (1.10) is a moment-generating function of a normal distribution, and I obtain the usual expression from the Vasicek (1977) model:

$$\ln B(t,T) = -\bar{r}^*\tau - (r_t - \bar{r}^*)\frac{1-e^{-b\tau}}{b} + \frac{\theta^2}{2b^2}\left(\tau - 2\frac{1-e^{-b\tau}}{b} + \frac{1-e^{-2b\tau}}{2b}\right), \tag{1.12}$$

where $\tau = T - t$ is the residual maturity.

Next, using the Feynman-Kac theorem[5], the forward price from t to T can be calculated as

$$\begin{aligned} G(t,T) &= \frac{\mathbb{E}_t^*\left[e^{-\int_t^T r_s \mathrm{d}s} S_T\right]}{B(t,T)} \\ &= \frac{1}{B(t,T)}\mathbb{E}_t^*\left[\exp\left(x_T - \int_t^T r_s \mathrm{d}s\right)\right]. \end{aligned} \tag{1.13}$$

The first equality, which also follows from a no-arbitrage argument, simply states that the discounted forward price from t to T should be the same as the expected discounted spot price at T. By integrating the equation (1.3), x_T can be expressed in the closed form as

$$x_T = \bar{x}^* - (\bar{x}^* - x_t)\, e^{-a\tau} + \sigma\int_t^T e^{-a(T-s)}\mathrm{d}W_s^* + \int_t^T e^{-a(T-s)}\ln U_s \mathrm{d}q_s^* - \lambda^* k^* \frac{1-e^{-a\tau}}{a}. \tag{1.14}$$

Using the assumption of independence between Wiener and Poisson processes, the expression (1.13) for the forward price can be written as

$$\begin{aligned} G(t,T) &= \frac{1}{B(t,T)}\mathbb{E}_t^*\left[\exp\left(x_T^0 - \int_t^T r_s \mathrm{d}s\right)\right] \\ &\quad \mathbb{E}_t^*\left[\exp\left(\int_t^T e^{-a(T-s)}\ln U_s \mathrm{d}q_s^* - \lambda^* k^* \frac{1-e^{-a\tau}}{a}\right)\right], \end{aligned} \tag{1.15}$$

where x_T^0 is the integral of equation (1.3) without the jump component ($k^* = \lambda^* = 0$):

$$x_T^0 = \bar{x}^* - (\bar{x}^* - x_t)\, e^{-a\tau} + \sigma\int_t^T e^{-a(T-s)}\mathrm{d}W_s^*. \tag{1.16}$$

[5] See, for example, Duffie (1992).

To calculate the last expectation in equation (1.15), note first that the integral in it can be written as

$$\int_t^T e^{-a(T-s)} \ln U_s \mathrm{d}q_s^* = \int_0^\tau e^{-a(\tau-s')} \ln U_{s'} \mathrm{d}q_{s'}^* = e^{-a\tau} \sum_{j=1}^{\nu_{t,T}} e^{as_j} \ln U_{s_j}, \quad (1.17)$$

where $s_j \in (0, \tau)$ are the moments in time (starting from t) when Poisson jumps occur (see Appendix B for the derivation). Under $\mathbb{P}^*$ the total number $\nu_{t,T}$ of jumps between t and T will be a Poisson variable with parameter $\lambda^*(T-t) = \lambda^*\tau$. Following Merton (1976), I will first assume that, for any t,

$$\ln U_t \, \Big| \, \nu_t = n \;\sim\; \mathcal{N}\left(n(\gamma_0^* - \omega^2/2), n\omega^2\right) \text{ i.i.d.} \quad n = 0, 1, \quad (1.18)$$

where $\nu_t \equiv \nu_{t,t+\mathrm{d}t}$. In other words, if a Poisson jump happens between t and $t + \mathrm{d}t$ (an event that happens with a probability $\lambda^*\mathrm{d}t$ under $\mathbb{P}^*$), then $\ln U_t$ will be a normal random variable with expectation $\gamma_0^* - \omega^2/2$ and variance ω^2. As $\varepsilon_t^*(U_t) = 1 + k^*$, it must be that $\gamma_0^* = \ln(1 + k^*)$. For finite time intervals, however, the corresponding distribution will depend on the fact that the underlying diffusion part of the process for S_t is GOU, rather than GBM as in Merton (1976). In the following I will assume that, given there are exactly n jumps in (t, T), the sum in (1.17) is also normally distributed under $\mathbb{P}^*$:

$$e^{-a\tau} \sum_{j=1}^{\nu_{t,T}} e^{as_j} \ln U_{s_j} \, \Big| \, \nu_{t,T} = n \;\sim\; \mathcal{N}\left(n(\gamma_\tau^* - \omega^2/2), n\omega^2\right) \text{ i.i.d.}, \quad (1.19)$$

where

$$\gamma_\tau^* = \ln(1 + k^*). \quad (1.20)$$

Specification (1.18) is then the instantaneous equivalent of (1.19), obtained if we let $\tau \to \mathrm{d}t$. Using (1.19), we obtain (see Appendix B for the derivation):

$$\begin{aligned} \ln G(t,T) &= \bar{x}^* - (\bar{x}^* - x_t)\, e^{-a\tau} + \frac{1}{2}\sigma^2 \frac{1 - e^{-2a\tau}}{2a} \\ &\quad - \frac{\rho\sigma\theta}{b}\left[\frac{1 - e^{-a\tau}}{a} - \frac{1 - e^{-(a+b)\tau}}{a+b}\right] + \lambda^* k^* \left(\tau - \frac{1 - e^{-a\tau}}{a}\right). \end{aligned}$$

$$(1.21)$$

Note that the terms proportional to θ^2 in the numerator and the denominator of equation (1.15) cancel out.

Price of a futures contract at time t with expiry at T can be also obtained from the Feynman-Kac theorem:

$$F(t,T) = \mathbb{E}_t^*\left[S_T\right]. \tag{1.22}$$

Intuitively, in the absence of arbitrage the futures price today on a contract expiring at T should be the same as today's expectation of the spot price at T. The basic difference between a forward and a futures contract is that they have different timing of cash flows. Although the net cash flows over the life of the contract are the same, in a forward contract money is exchanged only at the settlement date, whereas in a futures contract there is an exchange on a daily basis between the two parties through a system of margin calls. Also, in a (theoretical) forward contract, two counter-parties ex-ante assume symmetric risk; in a futures contract the counter-party risk is entirely taken by the exchange. The forward price is obtained by equating its present value with the present value of the spot contract at the settlement date. Cox, Ingersoll, and Ross (1981) prove that in general the two prices will be equal if the interest rates are deterministic. This clearly holds in the proposed model: if r_s is constant the integral in equation (1.15) is deterministic and cancels out with $B(t,T)$. Using this fact, by setting $\theta = 0$ in equation (1.21) we directly obtain

$$\ln F(t,T) = \bar{x}^* - \left(\bar{x}^* - x_t\right)e^{-a\tau} + \frac{1}{2}\sigma^2\frac{1-e^{-2a\tau}}{2a} + \lambda^* k^*\left(\tau - \frac{1-e^{-a\tau}}{a}\right). \tag{1.23}$$

With a slight abuse of notation, it is easy to verify that $F(t,T) \equiv F(S_t;t,T)$ given by equation (1.23) solves the following partial differential equation

$$\frac{\partial F(S_t;t,T)}{\partial t} + \left[a\left(m^* - \ln S_t\right) - \lambda^* k^*\right]S_t\frac{\partial F(S_t;t,T)}{\partial S_t} + \frac{1}{2}\sigma^2 S_t^2\frac{\partial^2 F(S_t;t,T)}{\partial S_t^2}$$

$$+\lambda^* \varepsilon_t^* \left[F(S_t U_t; t, T) - F(S_t; t, T) \right] = 0$$

with terminal boundary condition $F(S_T; T, T) = S_T$.

1.3.2 Implied Convenience Yield

Unlike financial assets, commodities are more than investment goods. They are often used for consumption or for industrial processes. This is why the notion of convenience yield is introduced for commodities. Convenience yield is the premium associated with holding the physical commodity rather than a contingent claim on it. The holding of an underlying good may become more profitable when market movements are irregular. For example, this happens if a good is in a short supply at a given moment. Then its relative scarcity drives the short term prices up, a phenomenon often called the ”inverted market”.

I do not use the convenience yield as an explicit pricing factor. There are two main reasons for this. Firstly, all the relevant price dynamics can be captured through the modeling of spot prices alone, and secondly, convenience yield is a theoretical construct introduced in order to explain backwardation in commodity futures. For any practical purposes it is more suitable to develop a model which works only with factors that are traded. However, the convenience yield is implicitly contained in the model. To derive the expression for the convenience yield, note that in the cost of carry model the following relation must hold:

$$CC(t, T) = y(t, T) - \Delta(t, T),$$

where $CC(t, T)$ is the spot cost-of-carry rate at t implied by a futures contract maturing at T, $y(t, T)$ is the yield at t on a unit discount bond maturing at T, while $\Delta(t, T)$ is the spot convenience yield implied by the futures contract (see Miltersen (2003)). The instantaneous implied forward convenience yield at t for a contract maturing at T can be calculated via relation

$$\delta(t,T) = \frac{\partial\left[(T-t)\Delta(t,T)\right]}{\partial T}. \tag{1.24}$$

Since

$$CC(t,T) = \frac{1}{T-t}\ln\left[\frac{F(t,T)}{S_t}\right],$$

we have

$$\begin{aligned}
\delta(t,T) &= -\frac{\partial\left[\ln B(t,T) + \ln F(t,T) - x_t\right]}{\partial T} \\
&= \bar{r}^* - \left(\bar{r}^* - r_t\right)e^{-b\tau} - \frac{1}{2}\theta^2\left(\frac{1-e^{-b\tau}}{b}\right)^2 \\
&\quad -a\left(\bar{x}^* - x_t\right)e^{-a\tau} - \frac{1}{2}\sigma^2 e^{-2a\tau} - \lambda^* k^*\left[a\tau - \left(1-e^{-a\tau}\right)\right]
\end{aligned} \tag{1.25}$$

As $t \to \infty$, the spot price and the interest rate converge in expectation to their long run equilibrium levels, $\bar{x}^*$ and $\bar{r}^*$, respectively. But then $\delta(t,T)$ will converge to $\bar{r}^* - \theta^2/2b^2$. This implies that the convenience yield also exhibits mean reversion, which is a feature of many models that use explicit dynamics for $\delta(t,T)$.

I can express $\delta(t,T)$ through the values of the bond price and the futures price only. I first use equations (1.12) and (1.23) to express r_t and x_t, respectively, through $\ln B(t,T)$ and $\ln F(t,T)$:

$$r_t = \bar{r}^* + \frac{b}{1-e^{-b\tau}}\left[\ln B(t,T) + \bar{r}^*\tau - \frac{\theta^2}{2b^2}\left(\tau - 2\frac{1-e^{-b\tau}}{b} + \frac{1-e^{-2b\tau}}{2b}\right)\right], \tag{1.26}$$

$$x_t = \bar{x}^* + e^{a\tau}\left[\ln F(t,T) - \bar{x}^* - \frac{1}{2}\sigma^2\frac{1-e^{-2a\tau}}{2a} - \lambda^* k^*\left(\tau - \frac{1-e^{-a\tau}}{a}\right)\right]. \tag{1.27}$$

Then, substituting these expressions into equation (1.25), we get $\delta(t,T)$ in terms of observable variables, $\ln F(t,T)$ and $\ln B(t,T)$.

The intuition behind results given by equation (1.25) is the following. Suppose that the term structure is flat ($r_t \equiv r$, $b = \theta = 0$). Then,

$$F(t,T) = S_t e^{(r-\delta)(T-t)}.$$

If commodity is held mostly for investment purposes, δ would be smaller than r. In this case $F(t,T) > S_t$, so it is more profitable to have a long position in a futures contract than hold the commodity itself. The described situation is usually known as "contango". On the other hand, if commodity is held primarily for consumption, the convenience yield is higher than the risk-free rate. In some situations it can be greater than r, so that $F(t,T) < S_t$. Thus, it may be better to own the physical commodity than the futures contract. This is usually referred to as "backwardation". Backwardation is a usual trademark of, say, energy commodities or industrial metals. For example, the market data indicate that crude oil is most of the time in backwardation, which is a consequence of typically high risk of supply shortage, moderately high transportation and storage costs, high consumption levels, and low value as a collateral for borrowing. In contrast, contango is typically observable in futures on investment commodities, such as gold or silver. These commodities have very low risk of being in short supply, very low transportation costs, low consumption levels relative to inventory, no risk of spoilage or loss whatsoever, no seasonality patterns in consumption or production, and finally they have very high value as a collateral.

When the futures curve is plotted against times to delivery, most of the commodities have a relatively stable long end and a rapidly changing short end. This is because the long end is more closely related to marginal costs of production, whereas the short end is governed by short-term supply and demand dynamics. Short end of the curve is usually used to cover unanticipated demand.

1.3.3 European Options on Futures

The value at time 0 of a European call option with strike price K, expiring at t, on a futures contract expiring at T is given by

$$C(0,t,T) = \mathbb{E}_0^* \left[e^{-\int_0^t r_s \mathrm{d}s} \max \left\{ F(t,T) - K, 0 \right\} \right]. \tag{1.28}$$

Denoting by $\nu_{0,t}$ the number of Poisson jumps occurring within the time interval $(0,t)$ and applying the rule for conditional expectation, we get

$$\begin{aligned} C(0,t,T) &= \sum_{n=0}^{\infty} \mathbb{E}_0^* \left[e^{-\int_0^t r_s \mathrm{d}s} \max \left\{ F(t,T) - K, 0 \right\} \Big| \nu_{0,t} = n \right] \mathbb{P}^* \left(\nu_{0,t} = n \right) \\ &= \sum_{n=0}^{\infty} e^{-\lambda^* t} \frac{(\lambda^* t)^n}{n!} \mathbb{E}_0^* \left[e^{-\int_0^t r_s \mathrm{d}s} \max \left\{ F(t,T) - K, 0 \right\} \Big| \nu_{0,t} = n \right] \end{aligned} \tag{1.29}$$

The presence of Poisson jumps causes the distribution of the spot price to be more skewed and leptokurtic than the log-normal. Option values are thus markedly changed by the jumps, since they are heavily influenced by one tail of the distribution.

Under the assumption (1.19) of log-normally distributed sizes of Poisson jumps U_t, the call price will be given by a series expansion of the form:

$$\begin{aligned} &C(0,t,T) = \\ &B(0,t) \sum_{n=0}^{\infty} e^{-\lambda^* t} \frac{(\lambda^* t)^n}{n!} \left\{ F(0,T) H(t,T) e^{\varphi(n,t,T)} N\left[d_1(n,\tau)\right] - K N\left[d_2(n,\tau)\right] \right\}, \end{aligned} \tag{1.30}$$

where $N(\cdot)$ is the probability function of a standard normal distribution,

$$H(t,T) = \exp \left\{ -\rho\sigma\theta \frac{e^{-a\tau}}{b} \left[\frac{1-e^{-at}}{a} - \frac{1-e^{-(a+b)t}}{a+b} \right] \right\}, \tag{1.31}$$

$$\begin{aligned} \varphi(n,t,T) &= n \left[e^{-a\tau} \left(\gamma_\tau^* - \frac{1}{2}\omega^2 \right) + e^{-2a\tau} \frac{1}{2}\omega^2 \right] \\ &\quad - \lambda^* k^* e^{-aT} \left[\frac{T}{1-e^{-aT}} - \frac{t}{1-e^{-at}} + \frac{e^{-at}}{a(1-e^{-at})} \right], \end{aligned} \tag{1.32}$$

$$
\begin{aligned}
d_1(n,\tau) &= \frac{\ln\left[F(0,T)H(t,T)/K\right] + \varphi(n,t,T) + v(n,\tau)/2}{\sqrt{v(n,\tau)}}, &(1.33)\\
d_2(n,\tau) &= d_1(n,\tau) - \sqrt{v(n,\tau)}, &(1.34)\\
v(n,\tau) &= e^{-2a\tau}\left(\sigma^2 \frac{1-e^{-2a\tau}}{2a} + n\omega^2\right). &(1.35)
\end{aligned}
$$

The proof is given in Appendix C. Although solution for the call price given by equation (1.30) is not a closed-form one, it has the form of a simple series expansion. Analogous result were obtained by Merton (1976) and Hilliard and Reis (1998) for the price of European call option on stock and commodity futures, respectively, when the underlying security follows a GBM with Poisson jumps. equation (1.32) implies that, as n increases, $\varphi(n,t,T)$ cannot grow faster than $n\left[e^{-a\tau}\left(\gamma_\tau^* - \omega^2/2\right) + e^{-2a\tau}\omega^2/2\right]$. This, in turn, means that all the terms in the sum on the right hand side of equation (1.30) are of the form $z^n/n!$, where z is a positive number. Given that $N(\cdot)$ is bounded, the series is clearly convergent and for any practical purposes it can be approximated by a finite sum.

When commodity spot prices follow GBM ($a \to 0$, $\bar{x}^* \to \infty$, and $a\bar{x}^* \to$ *const.*) and interest rates are deterministic ($\theta = 0$), equation (1.30) reduces to the futures call formula of Bates (1991). On the other hand, when jump component is absent ($k^* = \lambda^* = \omega = 0$) I recover the result of Miltersen and Scwartz (1998).

Formula for the price of a European put option on futures, $P(0,t,T)$, can be obtained by analogy. It is straightforward to show that

$$
\begin{aligned}
&P(0,t,T) = &(1.36)\\
&B(0,t)\sum_{n=0}^{\infty} e^{-\lambda^* t}\frac{(\lambda^* t)^n}{n!}\left\{KN\left[-d_2(n,\tau)\right] - F(0,T)H(t,T)e^{\varphi(n,t,T)}N\left[-d_1(n,\tau)\right]\right\}.
\end{aligned}
$$

1.3.4 Futures Hedge

Consider the minimum-variance hedge of a risk-averse investor. Duffie (1989) shows that optimal hedge ratio for an investor with mean-variance utility can be broken into two of the following portions: one reflecting speculative demand (which varies across individuals according to their risk aversion) and another reflecting a pure hedge (which is the same for all mean-variance utility hedgers). Because the former is both difficult to estimate and often close to zero, Duffie argues that it is reasonable to focus attention on the pure hedge.

Without loss of generality, assume that an investor at time t has taken a long position in a single spot contract and a short position in h_t futures contracts. The change in the worth of such a portfolio from t to $t + \mathrm{d}t$, under the risk-neutral measure $\mathbb{P}^*$, will be equal to

$$\mathrm{d}P_t^* = \mathrm{d}S_t^* - h_t \mathrm{d}F_t^*. \tag{1.37}$$

By taking the $\mathbb{P}^*$-variance of both sides of equation (1.37) conditional on the information available at t, we obtain

$$\mathrm{var}_t^*\left(\mathrm{d}P_t^*\right) = \mathrm{var}_t^*\left(\mathrm{d}S_t^*\right) - 2h_t \mathrm{cov}_t^*\left(\mathrm{d}S_t^*, \mathrm{d}F_t^*\right) + h_t^2 \mathrm{var}_t^*\left(\mathrm{d}F_t^*\right). \tag{1.38}$$

This expression is minimized by choosing a hedge ratio of

$$h_t^\bigstar = \frac{\mathrm{cov}_t^*\left(\mathrm{d}S_t^*,\ \mathrm{d}F_t^*\right)}{\mathrm{var}_t^*\left(\mathrm{d}F_t^*\right)}. \tag{1.39}$$

To compute $h_t^\bigstar$ one first needs to find the stochastic differential equation for $\mathrm{d}F_t^*$. To do so, I apply Itô's lemma to $F(t,T)$ obtained from equation (1.23) and get

$$\frac{\mathrm{d}F^*}{F} = -\lambda^* k^* \left(1 - e^{-a\tau}\right) \mathrm{d}t + e^{-a\tau}\left[\sigma \mathrm{d}W_t^* + (U_t - 1)\mathrm{d}q_t^*\right]. \tag{1.40}$$

The instantaneous optimal hedging ratio for dynamic futures hedge is thus

$$h_t^{\star} = \frac{S_t}{F(t,T)} e^{a\tau}. \tag{1.41}$$

Naturally, this ratio is truly optimal only for the interval $(t, t+\mathrm{d}t)$, after which it has to be updated to take into account the corresponding price changes.

I can easily generalize expression (1.41) to find an optimal hedging ratio for any calendar-spread hedge. If I assume that an investor at time t has taken a long position in one futures contract with maturity T_1 and a short position in h_t futures contracts with maturity $T_2 > T_1$, the optimal hedging ratio would be given by

$$h_t^{\star\ \text{cal.}} = \frac{F(t,T_1)e^{-aT_1}}{F(t,T_2)e^{-aT_2}}.$$

1.4 Calibration of the Model

1.4.1 Calibration Procedure

The empirical implementation of our model requires time-series observations for the state variables, namely the spot price and the short interest rate. In practice, these are seldom directly observable. However, the model can be casted in the state space form and the Kalman filter may be applied to estimate the parameters. The original Kalman filter approach is based on a supposition of normally distributed, serially uncorrelated, disturbances in the state variable. The model for the spot price of commodity developed in this chapter does not satisfy this assumption since the Poisson process distorts the distribution of return innovations. However, as shown in Harvey (1994), even if the disturbances belong to the family of affine jump diffusion, the Kalman filter recursion can still be applied under some approximations. The problem is that estimations with Kalman filter typically require a lot of computational time. Duan (1994) shows that when the transformation from unobservable to observable state variables is on an element-to-element basis, a simple maximum likelihood estimation can be applied. Namely, if for every t we can write $Y_t = \mathcal{F}_t(X_t, \psi)$, where $\mathcal{F}_t$ is a one-to-one mapping and ψ is the set of unknown parameters, then we can relate the likelihood function of observed variables to the likelihood function of the unobserved ones and the Jacobian matrix of the transformation:

$$\ln L_{\text{obs}}(\mathbf{Y}, \psi) = \ln L_{\text{un}}\left(\widetilde{\mathbf{X}}(\psi), \psi\right) + \ln\left|\det\left\{\left[\frac{\partial \mathcal{F}_t\left(\widetilde{\mathbf{X}}(\psi), \psi\right)}{\partial \mathbf{X}}\right]^{-1}\right\}\right|, \quad (1.42)$$

where $\mathbf{Y} = \{Y_t\}_{t=1}^N$, $\mathbf{X} = \{X_t\}_{t=1}^N$, and $\widetilde{X_t}(\psi) = \mathcal{F}_t^{-1}(Y_t, \psi)$ for every $t = 1, 2, \ldots, N$ (N being the length of the time series).

The general state space form applies to multivariate time series of observable variables (in this case, futures prices and bond prices for various maturities) related to unobservable ones (in this case spot prices[6] and short interest rates) via measurement equation. The measurement equation can be obtained from equations (1.23) and (1.12) by rewriting them in a discrete-time vector form:

$$Y_t = \alpha_t + \beta_t X_t, \quad (1.43)$$

where

$$\begin{aligned}
Y_t &= \begin{bmatrix} \ln F(t,T) \\ y(t,T) \end{bmatrix}, \\
X_t &= \begin{bmatrix} x_t \\ r_t \end{bmatrix}, \\
\alpha_t &= \begin{bmatrix} \bar{x}^*\left(1 - e^{-a\tau}\right) + \frac{1}{2}\sigma^2\frac{1-e^{-2a\tau}}{2a} + \lambda^* k^*\left(\tau - \frac{1-e^{-a\tau}}{a}\right) \\ \bar{r}^*\left(1 - \frac{1-e^{-b\tau}}{b\tau}\right) - \frac{\theta^2}{2b}\left(1 - 2\frac{1-e^{-b\tau}}{b\tau} + \frac{1-e^{-2b\tau}}{2b\tau}\right) \end{bmatrix},
\end{aligned}$$

[6]Spot contracts are seldom traded in a standardized form on exchanges. Exceptions, however, do exist when transportation costs are negligible. For example, Dubai Crude standard is regularly traded both on the spot and on the futures markets. Otherwise, the price of futures contract closest to maturity is often quoted as the "spot price".

$$\beta_t = \begin{bmatrix} e^{-a\tau} & 0 \\ 0 & \frac{1-e^{-b\tau}}{b\tau} \end{bmatrix}.$$

Here, I used the continuously compounded yield $y(t,T) = -[\ln B(t,T)]/(T-t)$ rather than the bond price $B(t,T)$.

The unobservable state variables are generated via the transition equation, which is simply a discrete-time version of the stochastic processes for the state variables under the physical probability $\mathbb{P}$, equations (1.9) and (1.2):

$$X_t = c + dX_{t-1} + \epsilon_t. \tag{1.44}$$

Here,

$$\begin{aligned} c \;&\equiv\; c(\nu_t) = \begin{bmatrix} (a\bar{x} - \lambda k)\,\frac{1-e^{-a\Delta t}}{a} \\ \bar{r}\left(1 - e^{-b\Delta t}\right) \end{bmatrix} + \nu_t \begin{bmatrix} \gamma_0 - \omega^2/2 \\ 0 \end{bmatrix}, \\ d \;&=\; \begin{bmatrix} e^{-a\Delta t} & 0 \\ 0 & e^{-b\Delta t} \end{bmatrix}, \end{aligned}$$

where $\nu_t \in \{0, 1\}$, $\gamma_0 = \ln(1+k)$, and Δt is the time interval between consecutive observations (in years). The error term ϵ_t is normally distributed under $\mathbb{P}$, conditionally on knowing whether the Poisson event occurs at t (i.e., $\nu_t = 1$) or not ($\nu_t = 0$):

$$\epsilon_t|\nu_t = n \;\sim\; \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix},\ \mathbf{V}(n)\right), \quad n = 0, 1, \tag{1.45}$$

where

$$\mathbf{V}(n) = \begin{bmatrix} \sigma^2\frac{1-e^{-2a\Delta t}}{2a} + n\omega^2 & \rho\sigma\theta\frac{1-e^{-(a+b)\Delta t}}{a+b} \\ \rho\sigma\theta\frac{1-e^{-(a+b)\Delta t}}{a+b} & \theta^2\frac{1-e^{-2b\Delta t}}{2b} \end{bmatrix}.$$

Therefore, conditionally on knowing ν_t, X_t will follow a VAR(1) process. Note that transition equation is expressed through parameters $\bar{x}$, $\bar{r}$, λ, and k, rather than the ones adjusted for the market price of risk. The reason is that equation (1.44) describes the joint process for commodity spot price and the interest rate under the physical (i.e. data-generating) measure $\mathbb{P}$. As we cannot observe x_t and r_t directly, we have to estimate $\bar{x}$, $\bar{r}$, λ, and k indirectly through their risk-adjusted counterparts $\bar{x}^*$, $\bar{r}^*$, λ^*, and k^*. Their relationship is given by equations (1.5)–(1.8). I assume that ξ_1 through ξ_4 are constant, although more generally they should depend on the business cycle and be correlated with the level of inventories. In summary, the set of parameters to be determined by estimation is

$$\psi = \{a,\ b,\ \bar{x}^*,\ \bar{r}^*,\ \lambda,\ k,\ \lambda^*,\ k^*,\ \sigma,\ \theta,\ \rho,\ \omega,\ \xi_1,\ \xi_2\}.$$

Under Bernoulli approximation for the Poisson jumps (that is, assuming $\nu_t \in \{0, 1\}$), the log-likelihood function for the unobserved data can be written as

$$\ln L_{\text{un}}\left(\mathbf{X}(\psi), \psi\right) = \tag{1.46}$$
$$\sum_{t=1}^{N-1} \ln\left[e^{-\lambda \Delta t} f\left(X_{t+1}, \psi | X_t, \nu_{t+1} = 0\right) + \left(1 - e^{-\lambda \Delta t}\right) f\left(X_{t+1}, \psi | X_t, \nu_{t+1} = 1\right)\right],$$

where $f(\cdot)$ is a conditional probability density function. Since x_t is normally distributed conditionally on knowing the exact number of Poisson jumps at $t + 1$, and r_t is unconditionally normal, the conditional probability density functions will be given by

$$f\left(X_{t+1}, \psi | X_t, \nu_{t+1} = n\right) =$$
$$\frac{1}{\sqrt{2\pi \ \det \mathbf{V}(n)}} \exp\left\{-\frac{1}{2}\left[X_{t+1} - \mu_{t+1}(n)\right]' \left[\mathbf{V}(n)\right]^{-1} \left[X_{t+1} - \mu_{t+1}(n)\right]\right\},$$

where

$$\mu_{t+1}(n) = \mathbb{E}_t(X_{t+1} | \nu_{t+1} = n) \ = \ c(n) + dX_t.$$

Applying the result given by equation (1.42), we can write the log-likelihood function for the observed data as

$$\ln L_{\text{obs}}\left(\mathbf{Y}(\psi),\psi\right) = \tag{1.47}$$
$$\sum_{t=1}^{N-1} \ln\left[e^{-\lambda\Delta t} f\left(\widetilde{X}_{t+1},\psi|X_t,\nu_{t+1}=0\right) + \left(1-e^{-\lambda\Delta t}\right) f\left(\widetilde{X}_{t+1},\psi|X_t,\nu_{t+1}=1\right)\right]$$
$$+ a\sum_{t=1}^{N-1}\tau_t - \sum_{t=1}^{N-1}\ln\left(\frac{1-e^{-b\tau_t}}{b\tau_t}\right),$$

where τ_t is the residual maturity on the contract at time t, and

$$\widetilde{X}_t = \beta_t^{-1}\left(Y_t - \alpha_t\right).$$

Maximization of the right hand side of equation (1.47) then yields the set of estimates for the unknown parameters, $\widetilde{\psi}$.

1.4.2 Data

The futures data I used for calibration consisted of weekly observations for futures prices of six commodities: two energy goods (Brent crude oil and natural gas), one industrial metal (copper), one precious metal (gold), one agricultural good (wheat), and one meat commodity (frozen pork bellies). I summarize these data in Table 1.1. The interest rate data consisted of yields on 3-Month Treasury Bills. All time series were obtained from Thomson Financial's Datastream and covered the same period – from May 29, 1991 to May 31, 2006, with 784 observations. The data were sampled on Wednesdays.[7] The weekly sampling frequency is chosen in order to reduce spurious market microstructure distortions and avoid weekend and other seasonal day-of-the-week effects. Furthermore, a higher sampling frequency would make estimation of mean jump frequency and mean jump size more difficult and more subject to noises in the data.

[7] If Wednesday was a holiday, Thursday data was used.

Table 1.1: **Commodity Futures: Data Summary**

Commodity	Exchange	Unit	Contract size	Average price†
Brent Crude Oil	ICE	\$/ barrel	1,000	25.26 (0.46)
Natural Gas	NYMEX	\$/ mmBtu	10,000	3.56 (0.09)
Copper	COMEX	¢/ lb	25,000	103.88 (1.45)
Gold	COMEX	\$/ troy oz.	100	353.90 (2.39)
Wheat	CBOT	¢/ bushel	5,000	338.66 (2.34)
Pork Bellies	CME	¢/ lb	40,000	66.23 (0.73)

† Standard errors in parentheses.

Trading period: May 29, 1991 – May 31, 2006.
Abbreviations: ICE – International Commodity Exchange, London. NYMEX – New York Mercantile Exchange. COMEX – Commodity Exchange, New York. CBOT – Chicago Board of Trade.
CME – Chicago Mercantile Exchange.
mmBtu – millions of British thermal units.

For different commodities and different time periods distinct specific futures contracts had to be used. The reason is that any futures contract has a limited time window when it is traded. Since prices of futures contracts are typically stacked in overlapping time series, a continuous series that spans many different contracts has to be created. Several methods are common for generating such continuous series (Rougier (1992)). I have chosen to consistently keep track of tenors that are second-to-closest to maturity. Since the contracts have a fixed maturity date, the time to maturity changes as time progresses. This is shown in Figure 1.1. Time to maturity for all the contracts in the sample was thus ranging between 12 and 33 trading days, with a mean of 22.45 and a standard deviation of 6.38.

Figures 1.2 through 1.8 show the plots of the time series used for calibration. Figure 1.2 displays the yield on 3-Month Treasury Bills, while Figures 1.3–1.8 show futures prices for different commodities. Some features, such as mean reversion or appearance of sudden jumps and dips, are clearly visible.

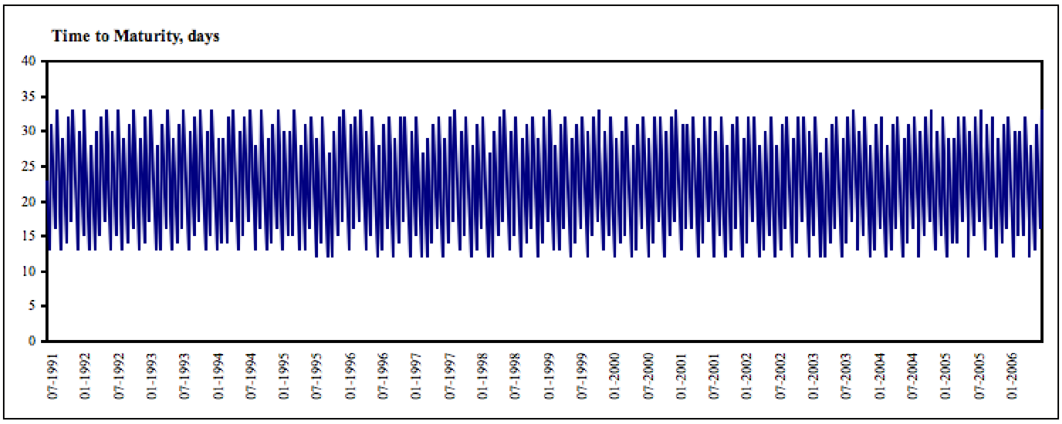

Figure 1.1: Time to maturity of the futures contracts used in the estimation.

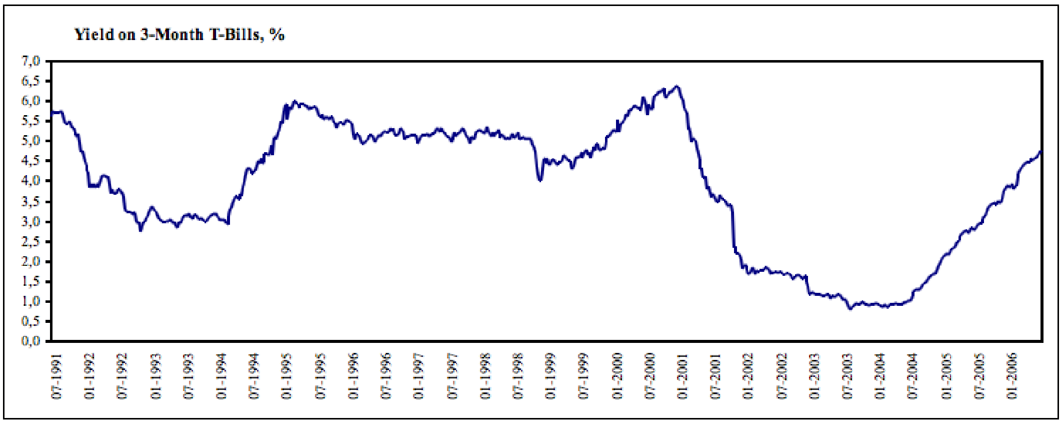

Figure 1.2: Yield on 3-month U.S. Treasury Bills, in percent.

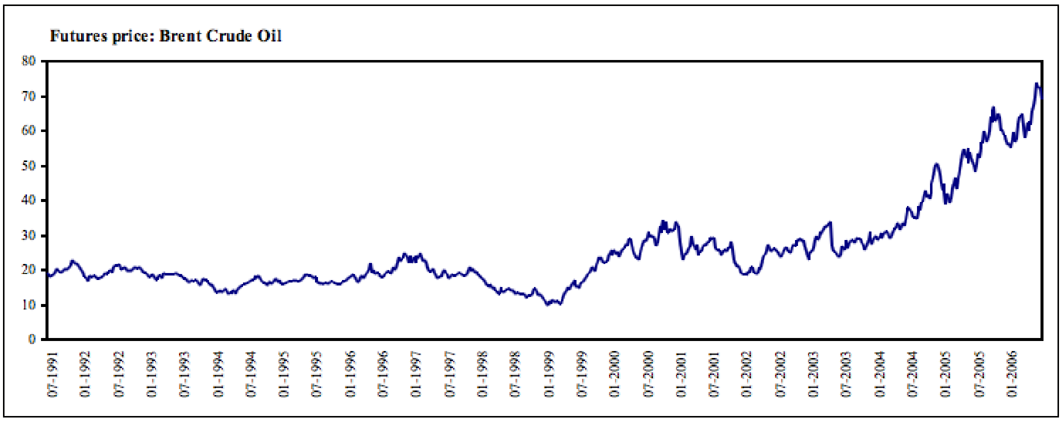

Figure 1.3: Price of the continuous series of futures contracts on Brent crude oil (USD per barrel).

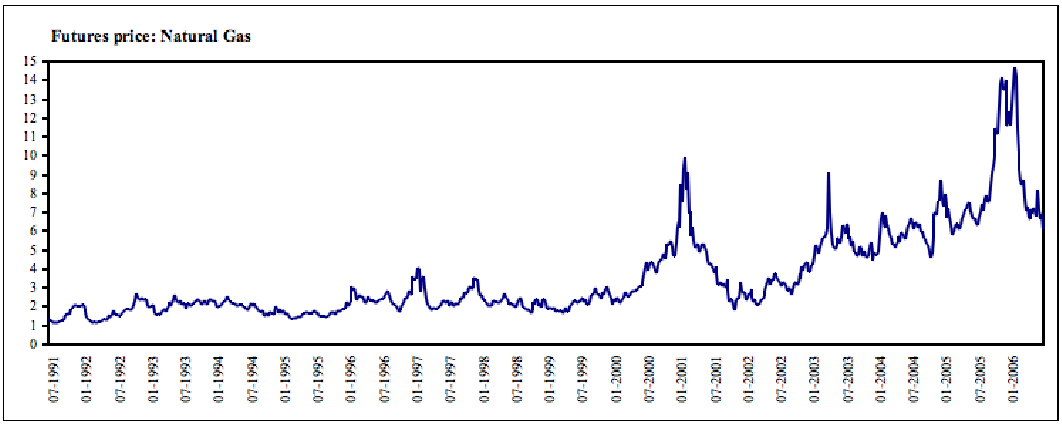

Figure 1.4: Price of the continuous series of futures contracts on natural gas (USD per million of Btu).

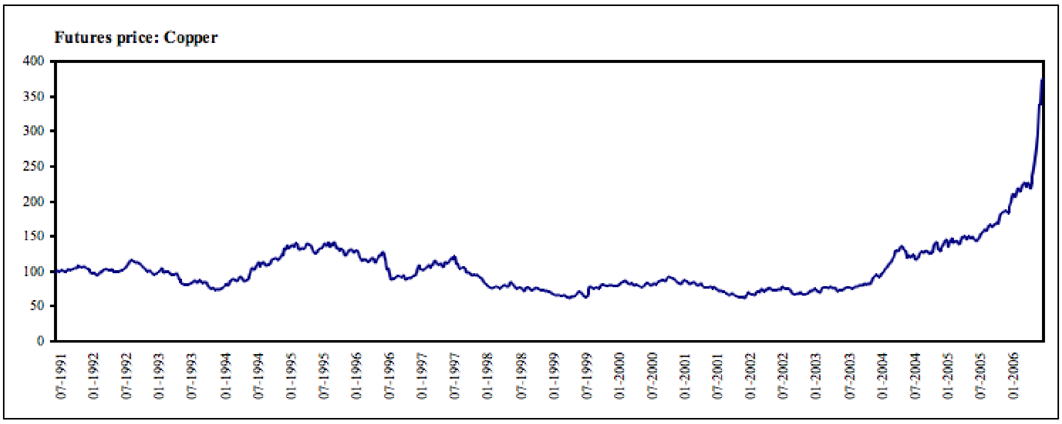

Figure 1.5: Price of the continuous series of futures contracts on copper (USD cents per pound).

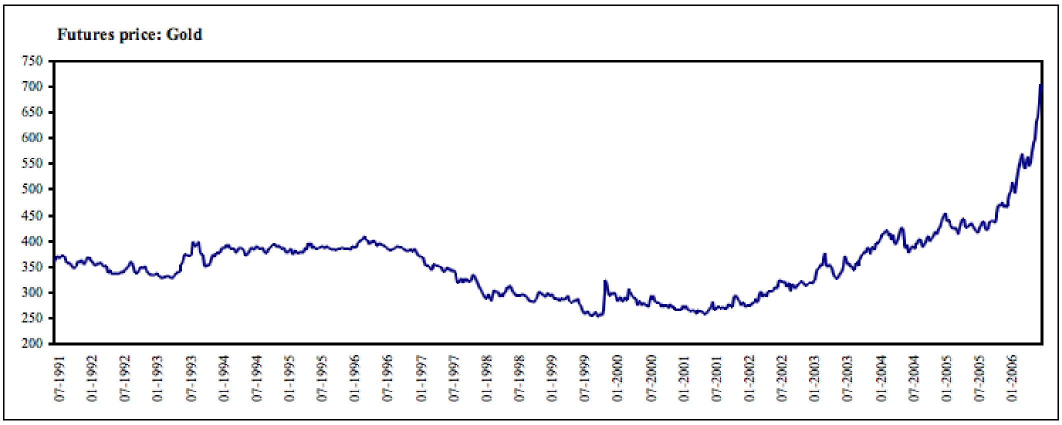

Figure 1.6: Price of the continuous series of futures contracts on gold (USD per troy ounce).

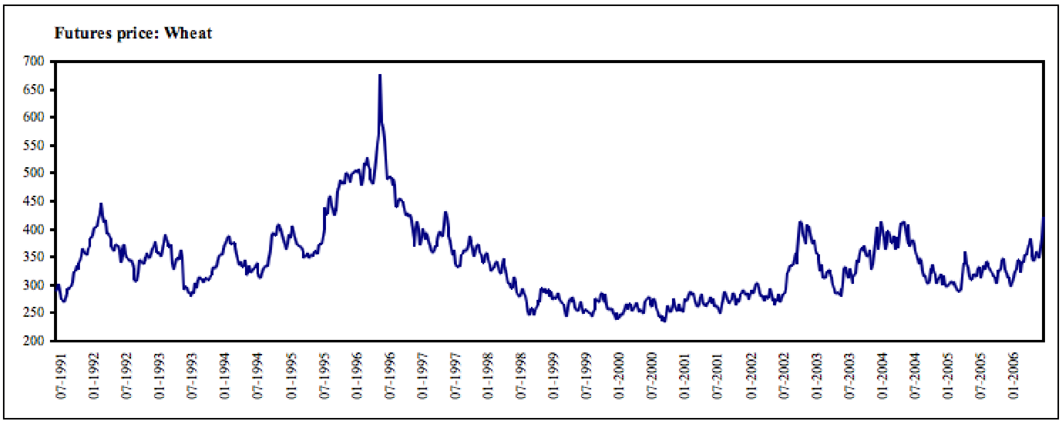

Figure 1.7: Price of the continuous series of futures contracts on wheat (USD cents per bushel).

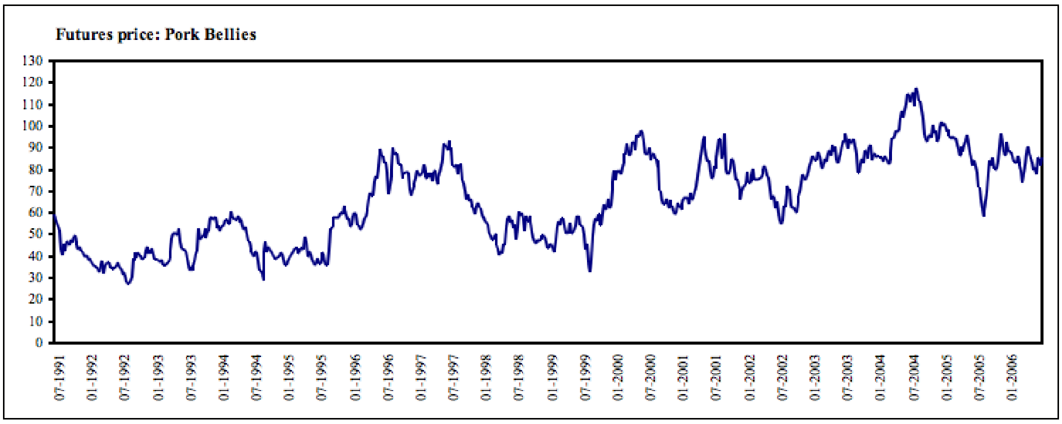

Figure 1.8: Price of the continuous series of futures contracts on pork bellies (USD cents per pound).

1.4.3 Results

The estimation is performed on inflation-adjusted deseazonalized time series of futures prices and the corresponding data on T-bills. I deflated each series of the futures prices with the CPI values obtained by interpolation of monthly data. Tables 1.2 through 1.7 summarize the results for the estimated parameters obtained by maximum likelihood technique described in Subsection 1.4.1. For each commodity, I give the comparison between the model and its two restrictions. Restriction 1 is a model without jumps ($\lambda = k = \omega = 0$), while restriction 2 also excludes the interest rate dependence ($b = \bar{r} = \theta = \rho = 0$). In the latter case the model reduces to the one-factor model of Schwartz (1997). For the baseline model and the two restrictions I report the values of log-likelihood function (LL) and Akaike information criterion (AIC) for each commodity.

Table 1.8 compares the likelihood ratios across the nested models. The critical values for χ^2 distribution at 1% confidence level when number of degrees of freedom is equal to 5, are given for comparison. (Here, $d = 5,\ 5$, and 10, respectively, count the degrees of freedom obtained when passing from more restricted model to a less restricted one.) Clearly, for all the commodities in the dataset both of the restrictions have to be rejected. In other words, a proper description of the commodity price dynamics has to include both the interest rate and the jumps. Also, in a two-factor model without the jump component Brownian diffusion alone cannot capture the whole dynamics of the spot price. As Akaike info criteria suggest, the two-factor model has similar explanatory power for all commodities in the sample.

As expected, all spot prices display significant mean reversion, which can be also seen in Figures 1.3 through 1.8. The fastest mean reversion is observed for natural gas and pork bellies: the characteristic times, measured as $1/a$, are around 31 weeks. Copper has the slowest mean reversion with characteristic time of about one year.

Another interesting result are small and typically negative correlations between interest rates and spot price returns of some commodities. (The values of ρ range from -0.0490 for gold to $+0.0035$ for wheat.) Similar numerical values were obtained by Schwartz (1997) with a three factor model for crude oil, gold, and copper. Economical reasons behind negative correlations are the following. High interest rates reduce the demand or increase the supply for storable commodities through a variety of channels. This may happen either due to increase of incentives for extraction of commodity today rather than tomorrow, or due to decrease in firms' desire to carry inventories (especially oil), or simply due to the fact that when interest rates increase the speculators are encouraged to shift out of their commodity contracts into treasury bills. These mechanisms work to reduce the market price of commodities, as indeed happened in the early 1980s. A decrease in real interest rates has the opposite effect, lowering the cost of carrying inventories and raising commodity prices, as was the case during 2001–2004 period.

Several interesting observations follow from the estimates of the jump parameters. For example, the frequency of jumps λ^* ranges between 0.1146 and 0.2684, or from about 6 to 14 abnormal weekly jumps per year. Apart from unexpected timing, jumps have also sizes that take a wide range of possible values. For example, natural gas has the largest standard deviation of jump sizes ($\omega = 0.0894$). The 95 percent confidence interval of $U - 1$ is therefore quite wide: roughly between -17 and $+18$ percent. The average jump size is positive ($k = 0.0094$), indicating that the long-run jump compensation is negative. On the other hand, gold has the narrowest jump size confidence interval, between -1.3 and $+1.8$ percent, with a positive average jump size ($k = 0.0027$).

Table 1.2: **Estimated Parameters: Brent Crude Oil**

Parameter	Model	Restriction 1	Restriction 2
a	1.5911 (0.6650)	1.3259 (0.6268)	1.1049 (0.4579)
b	2.0130 (0.2306)	2.3228 (0.1455)	-
$\bar{x}^*$	3.1947 (0.4319)	3.1989 (0.4933)	3.1964 (0.7915)
$\bar{r}^*$	0.0378 (0.0011)	0.0379 (0.0015)	-
λ^*	0.1556 (0.0114)	-	-
k^*	0.0015 (0.0010)	-	-
λ	0.2849 (0.0496)	-	-
k	0.0019 (0.0021)	-	-
σ	0.2986 (0.0677)	0.2488 (0.0944)	0.2073 (0.6934)
θ	0.0174 (0.0015)	0.0192 (0.0022)	-
ρ	-0.0263(0.0019)	-0.0219(0.0026)	-
ω	0.0208 (0.0226)	-	-
ξ_1	0.1613 (0.0894)	0.1346 (0.1013)	0.1490 (0.0800)
ξ_2	0.1186 (0.4512)	0.1192 (0.6191)	-
LL	6599.27	6589.52	2091.32
AIC	13170.53	13151.05	4154.63

(Standard errors in parentheses)

Table 1.3: **Estimated Parameters: Natural Gas**

Parameter	Model	Restriction 1	Restriction 2
a	1.6740 (0.3099)	1.1160 (0.1857)	1.1142 (0.1398)
b	1.0800 (0.4824)	2.1600 (0.7586)	-
$\bar{x}^*$	1.1619 (0.2868)	1.1387 (0.2983)	1.1243 (0.4137)
$\bar{r}^*$	0.0371 (0.0006)	0.0364 (0.0007)	-
λ^*	0.1146 (0.0043)	-	-
k^*	0.0094 (0.0031)	-	-
λ	0.2637 (0.0150)	-	-
k	0.0281 (0.0171)	-	-
σ	0.4509 (0.0882)	0.3006 (0.1215)	0.2802 (1.2950)
θ	0.0114 (0.0007)	0.0172 (0.0069)	-
ρ	-0.0055 (0.0007)	-0.0046 (0.0010)	-
ω	0.0894 (0.0224)	-	-
ξ_1	0.2134 (0.0229)	0.2166 (0.0513)	0.2707 (0.0622)
ξ_2	0.1624 (0.9470)	0.3248 (0.5882)	-
LL	6589.22	6572.84	1855.81
AIC	13150.43	13117.67	3683.63

(Standard errors in parentheses)

Table 1.4: **Estimated Parameters: Copper**

Parameter	Model	Restriction 1	Restriction 2
a	1.2742 (0.5122)	1.1584 (0.5340)	1.2268 (0.9231)
b	2.5245 (0.9121)	2.2950 (0.5178)	-
$\bar{x}^*$	4.6128 (0.2809)	4.6047 (0.2506)	4.6184 (0.3943)
$\bar{r}^*$	0.0355 (0.0041)	0.0352 (0.0041)	-
λ^*	0.2304 (0.0078)	-	-
k^*	0.0024 (0.0008)	-	-
λ	0.2902 (0.0249)	-	-
k	0.0036 (0.0035)	-	-
σ	0.2034 (0.1560)	0.1849 (0.1606)	0.1541 (0.0711)
θ	0.0211 (0.0092)	0.0192 (0.0089)	-
ρ	-0.0308 (0.0091)	-0.0385 (0.0092)	-
ω	0.0169 (0.0065)	-	-
ξ_1	0.4803 (0.1227)	0.5337 (0.1276)	0.5337 (0.5440)
ξ_2	0.4493 (0.2816)	0.4992 (0.3589)	-
LL	6595.51	6580.98	2312.65
AIC	13163.03	13133.97	4597.30

(Standard errors in parentheses)

Table 1.5: **Estimated Parameters: Gold**

Parameter	Model	Restriction 1	Restriction 2
a	1.1326 (0.2425)	1.0296 (0.8582)	0.8972 (1.0983)
b	2.5424 (1.8600)	2.3113 (1.2827)	-
$\bar{x}^*$	5.8889 (0.8369)	5.8890 (0.8646)	5.8738 (1.7478)
$\bar{r}^*$	0.0349 (0.0087)	0.0345 (0.0095)	-
λ^*	0.1620 (0.0117)	-	-
k^*	0.0027 (0.0019)	-	-
λ	0.2490 (0.0404)	-	-
k	0.0040 (0.0098)	-	-
σ	0.1170 (0.0144)	0.1064 (0.0165)	0.0887 (0.0705)
θ	0.0211 (0.0021)	0.0192 (0.0023)	-
ρ	-0.0490 (0.0196)	-0.0613 (0.0220)	-
ω	0.0080 (0.0060)	-	-
ξ_1	0.6220 (0.3610)	0.6126 (0.3743)	0.7657 (1.8183)
ξ_2	0.5513 (0.7912)	0.6126 (1.1261)	-
LL	6589.55	6575.90	2743.61
AIC	13151.10	13123.80	5459.21

(Standard errors in parentheses)

Table 1.6: **Estimated Parameters: Wheat**

Parameter	Model	Restriction 1	Restriction 2
a	0.9854 (0.4268)	0.8958 (0.2789)	0.7465 (0.1181)
b	2.5422 (1.6166)	2.3111 (1.5254)	-
$\bar{x}^*$	5.8908 (0.7310)	5.8899 (1.0358)	5.8905 (1.7239)
$\bar{r}^*$	0.0384 (0.0093)	0.0383 (0.0142)	-
λ^*	0.2684 (0.0176)	-	-
k^*	0.0005 (0.0006)	-	-
λ	0.3383 (0.0690)	-	-
k	0.0007 (0.0005)	-	-
σ	0.1212 (0.0130)	0.1102 (0.0204)	0.0918 (0.0711)
θ	0.0211 (0.0018)	0.0192 (0.0027)	-
ρ	0.0035 (0.0027)	0.0029 (0.0045)	-
ω	0.0092 (0.0061)	-	-
ξ_1	0.0703 (0.0375)	0.0781 (0.0638)	0.0728 (0.1998)
ξ_2	0.0524 (0.0327)	0.0582 (0.0660)	-
LL	6584.42	6571.47	2712.66
AIC	13140.84	13114.94	5397.31

(Standard errors in parentheses)

Table 1.7: **Estimated Parameters: Pork Bellies**

Parameter	Model	Restriction 1	Restriction 2
a	1.6433 (0.6073)	1.4939 (0.6202)	1.2449 (0.8914)
b	2.5553 (1.8095)	2.3230 (1.5197)	-
$\bar{x}^*$	4.2042 (0.4889)	4.2061 (0.6791)	4.2099 (0.5163)
$\bar{r}^*$	0.0378 (0.0092)	0.0376 (0.0144)	-
λ^*	0.2579 (0.0170)	-	-
k^*	0.0010 (0.0006)	-	-
λ	0.4054 (0.0103)	-	-
k	0.0012 (0.0008)	-	-
σ	0.2177 (0.0377)	0.1979 (0.0608)	0.1649 (0.1456)
θ	0.0211 (0.0017)	0.0192 (0.0025)	-
ρ	-0.0220 (0.0017)	-0.0183 (0.0027)	-
ω	0.0216 (0.0168)	-	-
ξ_1	0.2369 (0.1184)	0.2154 (0.1691)	0.1795 (0.3305)
ξ_2	0.1383 (0.0437)	0.1537 (0.0874)	-
LL	6612.60	6596.38	2276.85
AIC	13197.19	13164.76	4525.70

(Standard errors in parentheses)

Table 1.8: **Likelihood Ratios**

Commodity	Model / Restr. 1	Restr. 1 / Restr. 2
Brent Crude Oil	19.49	8996.41
Natural Gas	32.76	9434.05
Copper	29.06	8536.67
Gold	27.31	7664.58
Wheat	25.90	7717.63
Pork Bellies	32.43	8639.06
χ^2_5 $(p = 0.01)$	15.09	

Table 1.9 shows the values of three types of risk premia implied by the model (cf. Section 2). The values are stated in percentage points. As expected, investors are more compensated for being exposed to diffusive risk contained in the spot price movements than to the corresponding interest rate risk. The spot price diffusive-risk premium is highest for copper (7.45%), which is reasonable given its inelastic and slow responses to longer periods of increases or decreases in price. Diffusive risk explains more than 95 percent of the overall risk premium for crude oil, copper and gold. On the other hand, the interest rate diffusive risk pays substantially lower premia. This is mostly because the volatility of the interest rate is, in some cases, an order of magnitude lower than the volatility of the spot price: θ is estimated to be around 0.02, whereas the lowest volatility of the spot price is that of gold with $\sigma = 0.1170$.

Market prices also carry important information about investors' expectations of abnormal movements in the future. This is mirrored through very high jump-risk premia for some of the commodities in the sample. The values of these premia are comparable to the diffusive return risk. Of the commodities chosen, natural gas has the highest combined jump-time/jump-size premium of 1.74%, followed by pork bellies (1.49%). The highest value obtained for nat-

Table 1.9: **Risk Premia (percentage points)**

Commodity	Spot price diffusive risk	Interest rate diffusive risk	Jump risk (overall)
Brent Crude Oil	2.97	0.07	0.12
Natural Gas	5.72	0.22	1.74
Copper	7.45	0.06	0.25
Gold	6.15	0.05	0.12
Wheat	0.87	0.04	0.25
Pork Bellies	3.10	0.13	1.49

ural gas comes from the highest disproportion between the risk-neutral mean jump size (k^*) and the actual one (k) among all six commodities. This difference is directly related to the volatility of jump size (cf. Appendix A), which is highest for natural gas.

The significance of jumps indicates the presence of a systemic and non-diversifiable risk component. To explore the economic interpretation of this component, let us analyze what drives the combined jump premium. The combination of a high uncertainty of jump size and a high uncertainty of jump occurrence is reflected on the two components of the jump risks. Firstly, it increases the jump-time risk through a large difference between "physical" λ, and the one perceived by the market, λ^*, which is relevant for derivatives pricing. Secondly, it increases jump-size risk by driving a large difference between k and k^*. As the empirical results suggest, the differences between λ and λ^* are typically much greater than the differences between k and k^*. Therefore, most of the commodity risk premium can be attributed to compensation for investors' aversion towards rare events that cause the jumps. In a similar context of rare-event uncertainty embedded in stock options, Pan (2002) estimates that the premium demanded by equity investors for this type of risk is around 3.5% per year. Liu, Pan, and Wang (2005) attribute a part of the jump risk to the uncertainty aversion in the sense of Knight (1921) and Ellsberg (1961).

To complete the exposition of numerical results, in Tables 1.10 and 1.11 I report the values of European call and put options on futures, obtained using equations (1.30) and (1.36), respectively. The sum in these two equations converge to precision of 10^{-4} after only four or five terms. For each commodity I use the point estimates of parameters obtained by calibration. The strike prices chosen are arbitrary, the maturity of each option is $t = 150$ trading days, while the prices of futures contracts are those quoted on June 1, 2006. Futures contracts, which are different from the ones used to calibrate the parameters, were selected such that their delivery dates were as close as possible to March 1, 2007. Times to maturity are also shown in the table. Yield on the 3-Month T-Bills was 4.71%. I also report the premia to option prices due to Poisson jumps. These premia (when measured relative to the total price) are particularly high for deep in- or out-of-the-money options, but also for commodities with high expected jump frequencies and/or volatile jump sizes. Good examples are Brent and natural gas.

Table 1.10: **Prices of European call options on commodity futures**

Commodity	K	$F(0,T)$	T	$C(0,t,T)$	Jump premium
Brent Crude Oil	70.00	72.31	273	3.2164	0.5125
Natural Gas	10.50	10.14	270	0.2523	0.0459
Copper	310.00	326.85	273	18.1683	2.9954
Gold	650.00	652.70	270	10.7545	1.6507
Wheat	472.50	468.75	267	6.0518	1.1481
Pork Bellies	85.00	84.25	267	1.2901	0.2877

Table 1.11: **Prices of European put options on commodity futures**

Commodity	K	$F(0,T)$	T	$P(0,t,T)$	Jump premium
Brent Crude Oil	70.00	72.31	273	0.9657	0.1520
Natural Gas	10.50	10.14	270	0.5940	0.0837
Copper	310.00	326.85	273	1.7228	0.2666
Gold	650.00	652.70	270	7.9771	1.0037
Wheat	472.50	468.75	267	9.6699	1.7566
Pork Bellies	85.00	84.25	267	2.0139	0.4373

1.5 Conclusion

The chapter presents a simple two-factor model for valuation of contingent claims on commodities. The logarithm of the spot price is modeled by an Ornstein-Uhlenbeck process with compensated Poisson jumps. The interest rate is considered to follow the Vasicek process. The stochastic processes chosen are such that they are the minimal ones capturing the desired dynamics, yet providing closed form solutions. By transforming the factor dynamics in a way that the processes are expressed under the risk-neutral measure I have obtained the closed form solutions for the price of a forward and a future contract, and a formula for the price of a European option on commodity futures in form of a series expansion. Furthermore, I have derived an expression for the optimal hedging ratio of a dynamic futures hedge by applying the minimum-variance hedging. Full analytical tractability of the model allowed for parameter estimation via usual maximum likelihood technique, adapting it to account for jumps in the spot price. In this way I avoid computationally intensive Kalman filter procedure.

The principal advantages of the proposed model over the ones in the literature are the following. First, it takes into account the most important

Table 1.12: **European option prices on natural gas futures as a function of moneyness**

$K/F(0,T)$	$C(0,t,T)$	Jump premium	$P(0,t,T)$	Jump premium
0.7	0.2925	0.0440	0.0000	0.0000
0.8	0.1956	0.0298	0.0004	0.0001
0.9	0.1048	0.0166	0.0068	0.0011
1.0	0.0392	0.0068	0.0384	0.0055
1.1	0.0095	0.0019	0.1058	0.0148
1.2	0.0015	0.0004	0.1951	0.0275
1.3	0.0002	0.0001	0.2910	0.0414

stylized facts for prices of commodities and bonds. Second, it avoids explicit specification of dynamics of the convenience yield. Third, and perhaps most important, it allows for the closed form solution for basic contingent claims: forward and futures price, and price of a European option on futures. Finally, the calibration can be done in a "natural" fashion, i.e. it requires only the data for commodity futures and bond yields.

The estimations using the actual data indicate that, in addition to spot price diffusion, interest rate diffusion and jumps cannot be neglected when derivative contracts on commodities are priced. Furthermore, the values of market prices of risks implied by the model indicate that for some commodities jump risk pays very high premia. These values are comparable to the premia originating from the spot price diffusion, and they are certainly much greater than the ones originating from the interest rate diffusion.

Both the analytical results and parameter estimations, however, may be sensitive to the fact that the stochastic nature of the spot price volatility is neglected. Setting up and solving a more general model that allows for conditional heteroskedasticity in both the spot price and the interest rate would be a natural extension of the results presented in this chapter. In this way, the discrete-time approximation of the model will belong to the ARCH family and become applicable for more robust value-at-risk estimations. Additional generalization should allow for time-dependent long-run equilibrium levels. The justification for this generalization can be sought in the possibility of non-cointegrated supply and demand series, which is the case with some extracted commodities such as crude oil. While commodities such as industrial metals and cultivated agricultural goods have, to a fair extent, cointegrated supply and demand, similar observation is unlikely to hold for energy commodities where demand is growing while at the same time the global supply is gradually diminishing.

The results of this chapter can be also readily generalized to include numerical pricing techniques for contracts with more complex payoff functions. Examples of interest include American or exotic options on futures, multiperiod budgeting decisions, or dynamic optimization in the real-option setup.

Appendix A: Risk-neutral processes

To find the risk-neutral equivalents of equations (1.1) and (1.2), note that we can split the dynamics for the commodity spot price into

$$\frac{\mathrm{d}S_t}{S_t} = \left(\frac{\mathrm{d}S_t}{S_t}\right)_{\mathrm{diff}} + \mathrm{d}J_t, \tag{1.48}$$

where

$$\left(\frac{\mathrm{d}S_t}{S_t}\right)_{\mathrm{diff}} = a\left(m - \ln S_t\right)\mathrm{d}t + \sigma \mathrm{d}W_t \tag{1.49}$$

is the diffusion part, while

$$\mathrm{d}J_t = (U_t - 1)\mathrm{d}q_t - \lambda k \mathrm{d}t \tag{1.50}$$

is the jump part of the stochastic process followed by S_t. Let us focus on the diffusion part first. Equation (2.18), together with (1.2), follows a joint Brownian diffusion, since $\mathrm{d}W$ and $\mathrm{d}Z$ are correlated. Define

$$\mathrm{d}\mathbf{B_t} = \begin{bmatrix} \mathrm{d}W_t \\ \mathrm{d}Z_t \end{bmatrix}. \tag{1.51}$$

To find the risk-neutral equivalent $\mathrm{d}\mathbf{B}^*$ of (2.20), which would be a martingale under an equivalent measure $\mathbb{P}^*$, we first write the Radon-Nikodým derivative of $\mathbb{P}^*$ with respect to the physical measure $\mathbb{P}$:

$$L_t \equiv \frac{\mathrm{d}\mathbb{P}^*}{\mathrm{d}\mathbb{P}} = \exp\left[-\int_0^t \boldsymbol{\xi}_s \cdot \mathrm{d}\mathbf{B}_s - \frac{1}{2}\int_0^t \left(\boldsymbol{\xi}_s \cdot \mathrm{d}\mathbf{B}_s\right)\left(\mathrm{d}\mathbf{B}_s \cdot \boldsymbol{\xi}_s\right)\right],$$

where

$$\boldsymbol{\xi}_s = \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix}$$

is predictable at s (Bingham and Kiesel (2004)). Then, by Girsanov's theorem, a $\mathbb{P}^*$-Brownian motion has the form

$$\mathrm{d}\mathbf{B}_t^* = \mathrm{d}\mathbf{B}_t\left(1 + \mathrm{d}\mathbf{B}_t \cdot \boldsymbol{\xi}_t\right).$$

Therefore,

$$\begin{aligned} \mathrm{d}\mathbf{B}_t &= \mathrm{d}\mathbf{B}_t^* - \mathrm{d}\mathbf{B}_t\left(\mathrm{d}\mathbf{B}_t \cdot \boldsymbol{\xi}_t\right) \\ &= \mathrm{d}\mathbf{B}_t^* - \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} \mathrm{d}t, \end{aligned}$$

which implies

$$\begin{bmatrix} \mathrm{d}W_t \\ \mathrm{d}Z_t \end{bmatrix} = \begin{bmatrix} \mathrm{d}W_t^* - (\xi_1 + \rho\xi_2)\,\mathrm{d}t \\ \mathrm{d}Z_t^* - (\rho\xi_1 + \xi_2)\,\mathrm{d}t \end{bmatrix}$$

for all t. Hence, the processes

$$\left(\frac{\mathrm{d}S_t^*}{S_t}\right)_{\text{diff}} = a\left[m - \frac{\sigma}{a}\left(\xi_1 + \rho\xi_2\right) - \ln S_t\right]\mathrm{d}t + \sigma \mathrm{d}W_t^*$$

and

$$\mathrm{d}r_t^* = b\left[\bar{r} - \frac{\theta}{b}\left(\rho\xi_1 + \xi_2\right) - r_t\right]\mathrm{d}t + \sigma\mathrm{d}Z_t^*$$

contain diffusions that are martingales under $\mathbb{P}^*$.

On the other hand, the jump component, equation (2.19), is a $\mathbb{P}$-martingale, since

$$\begin{aligned}\mathbb{E}_t(\mathrm{d}J_t) &= \mathbb{E}_t\left[(U_t - 1)\mathrm{d}q_t\right] - \lambda k\mathrm{d}t \\ &= \varepsilon_t(U_t - 1)\lambda\mathrm{d}t - \lambda k\mathrm{d}t \\ &= 0.\end{aligned}$$

By applying Girsanov's theorem for point processes (Elliot and Kopp (2005)) and using the normality assumption (1.18), we find

$$\begin{aligned}\mathrm{d}q_t^* &= \mathrm{d}q_t - \varepsilon_t\left\{\exp\left[\xi_3 + \xi_4\ln U_t - \xi_4\left(\gamma_0 - \frac{1}{2}\omega^2\right) + \frac{1}{2}\xi_4^2\omega^2\right] - 1\right\}\lambda\mathrm{d}t \\ &= \mathrm{d}q_t - \left\{\varepsilon_t\left[\exp\left(\xi_4\ln U_t\right)\right]\exp\left[\xi_3 - \xi_4\left(\gamma_0 - \frac{1}{2}\omega^2\right) + \frac{1}{2}\xi_4^2\omega^2\right] - 1\right\}\lambda\mathrm{d}t \\ &= \mathrm{d}q_t - (e^{\xi_3} - 1)\lambda\mathrm{d}t.\end{aligned}$$

Define $\lambda^* = \lambda e^{\xi_3}$. Then,

$$\mathrm{d}q_t^* = \mathrm{d}q_t + \lambda\mathrm{d}t - \lambda^*\mathrm{d}t.$$

Girsanov's theorem applied to $\mathrm{d}J$ then yields

$$\begin{aligned}\mathrm{d}J_t^* &= \mathrm{d}J_t - \mathbb{E}_t\left[\left\{\exp\left[\xi_3 + \xi_4\ln U_t - \xi_4\left(\gamma_0 - \frac{1}{2}\omega^2\right) + \frac{1}{2}\xi_4^2\omega^2\right] - 1\right\}(U_t - 1)\mathrm{d}q_t\right] \\ &= \mathrm{d}J_t - \varepsilon_t\left[\left\{\exp\left[\xi_3 + \xi_4\ln U_t - \xi_4\left(\gamma_0 - \frac{1}{2}\omega^2\right) + \frac{1}{2}\xi_4^2\omega^2\right] - 1\right\}(e^{\ln U_t} - 1)\right]\lambda\mathrm{d}t \\ &= \mathrm{d}J_t - \left(e^{\xi_3+\xi_4\omega^2+\gamma_0} - e^{\xi_3} - e^{\gamma_0} + 1\right)\lambda\mathrm{d}t.\end{aligned}$$

Define $1 + k^* \equiv e^{\xi_4 \omega^2 + \gamma_0} = (1+k)e^{\xi_4 \omega^2}$. The process

$$\begin{aligned} \mathrm{d}J_t^* &= \mathrm{d}J_t - \left[(1+k^*)e^{\xi_3} - e^{\xi_3} - (1+k)\right] \lambda \mathrm{d}t \\ &= (U_t - 1)\mathrm{d}q_t - \lambda^* k^* \mathrm{d}t \end{aligned}$$

is then a martingale under $\mathbb{P}^*$, with $\mathbb{E}_t^*(\mathrm{d}J_t^*) = 0$. Putting everything together, the processes

$$\frac{\mathrm{d}S_t^*}{S_t} = a\left[m - \frac{\sigma}{a}(\xi_1 + \rho\xi_2) - \ln S_t\right]\mathrm{d}t + \sigma \mathrm{d}W_t^* + (U_t - 1)\mathrm{d}q_t - \lambda^* k^* \mathrm{d}t$$

and

$$\mathrm{d}r_t^* = b\left[\bar{r} - \frac{\theta}{b}(\rho\xi_1 + \xi_2) - r_t\right]\mathrm{d}t + \sigma \mathrm{d}Z_t^*$$

represent the risk-neutral equivalents of (1.1) and (1.2).

Appendix B: Price of a forward contract

To show that equation (1.21) is equivalent to equation (1.15), note first that the diffusion part of the log spot price,

$$x_T^0 = \bar{x}^* - (\bar{x}^* - x_t)\, e^{-a(T-t)} + \sigma \int_t^T e^{-a(T-s)} \mathrm{d}W_s^*, \tag{1.52}$$

is distributed normally under $\mathbb{P}^*$, with mean

$$\mathbb{E}_t^*\left(x_T^0\right) = \bar{x}^* - (\bar{x}^* - x_t)\, e^{-a(T-t)}$$

and variance

$$\begin{aligned} \mathrm{var}_t^*\left(x_T^0\right) &= \mathrm{var}_t^*\left[\sigma \int_t^T e^{-a(T-s)} \mathrm{d}W_s^*\right] \\ &= \sigma^2 \int_t^T e^{-2a(T-t)} \mathrm{d}s' \end{aligned}$$

$$= \sigma^2 \frac{1-e^{-2a(T-t)}}{2a}.$$

Here I have applied the isometry of Itô integral to calculate the variance. Similarly, we find

$$\begin{aligned}
\mathrm{cov}_t^* \left(x_T^0, \int_t^T r_s \mathrm{d}s \right) &= \sigma\theta \mathbb{E}_t^* \left[\int_t^T e^{-a(T-s)} \mathrm{d}W_s^* \int_t^T \frac{1-e^{-b(T-s)}}{b} \mathrm{d}Z_s^* \right] \\
&= \rho\sigma\theta \int_t^T e^{-a(T-s)} \frac{1-e^{-b(T-s)}}{b} \mathrm{d}s \\
&= \rho\sigma\theta \left[\frac{1-e^{-a(T-t)}}{a} - \frac{1-e^{-(a+b)(T-t)}}{a+b} \right].
\end{aligned}$$

Hence,

$$\begin{aligned}
\mathbb{E}_t^* \left[\exp\left(x_T^0 - \int_t^T r_s \mathrm{d}s \right) \right] &= \mathbb{E}_t^* \left[\exp\left(x_T^0 \right) \right] \mathbb{E}_t^* \left[\exp\left(- \int_t^T r_s \mathrm{d}s \right) \right] \\
&\quad \cdot \exp\left[\mathrm{cov}_t^* \left(x_T^0, - \int_t^T r_s \mathrm{d}s \right) \right] \\
&= \exp\left[\bar{x}^* - (\bar{x}^* - x_t)\, e^{-a\tau} + \frac{1}{2}\sigma^2 \frac{1-e^{-2a\tau}}{2a} \right] B(t,T) \\
&\quad \cdot \exp\left[-\rho\sigma\theta \left(\frac{1-e^{-a\tau}}{a} - \frac{1-e^{-(a+b)\tau}}{a+b} \right) \right]. \quad (1.53)
\end{aligned}$$

Next, to compute the expectation

$$\mathbb{E}_t^* \left[\exp\left(\int_t^T e^{-a(T-s)} \ln U_s \mathrm{d}q_s^* - \lambda^* k^* \frac{1-e^{-a\tau}}{a} \right) \right]$$

we use the fact that the integral over $\mathrm{d}q_s^*$ can be written as

$$\int_t^T e^{-a(T-s)} \ln U_s \mathrm{d}q_s^* = \sum_{j=1}^{\nu_{t,T}} \int_t^T e^{-a(T-s)} \ln U_s \delta(s-s_j) \mathrm{d}s = e^{-a\tau} \sum_{j=1}^{\nu_{t,T}} e^{a s_j} \ln U_{s_j},$$

where $\delta(s - s_j)$ is the Dirac distribution, i.e. a Riemann-integrable infinity at instants s_j when jumps occur, and zero otherwise. The jumps are counted by a discrete variable j, and the total number $\nu_{t,T}$ of jumps in (t,T) is a Poisson random variable under $\mathbb{P}^*$. Under the assumption of normality, equation (1.19), we find

$$
\begin{aligned}
&\mathbb{E}_t^*\left[\exp\left(\int_t^T e^{-a(T-s)}\ln U_s \mathrm{d}q_s^*\right)\right] = \mathbb{E}_t^*\left[\exp\left(e^{-a\tau}\sum_{j=1}^{\nu_{t,T}} e^{as_j}\ln U_{s_j}\right)\right] \\
= &\sum_{n=0}^{\infty}\varepsilon_t^*\left[\exp\left(e^{-a\tau}\sum_{j=1}^{\nu_{t,T}} e^{as_j}\ln U_{s_j}\right)\Big|\nu_{t,T}=n\right]\mathbb{P}^*(\nu_{t,T}=n) \\
= &\sum_{n=0}^{\infty}\exp\left[n(\gamma_\tau^*-\omega^2/2)+n\omega^2/2\right]\mathbb{P}^*(\nu_{t,T}=n) \\
= &\sum_{n=0}^{\infty}\exp\left(n\gamma_\tau^*\right)\mathbb{P}^*(\nu_{t,T}=n) \\
= &\sum_{n=0}^{\infty}\left(1+k^*\frac{e^{-a\tau}}{1-e^{-a\tau}}\right)^n e^{-\lambda^*\tau}\frac{(\lambda^*\tau)^n}{n!} \\
= &e^{\lambda^* k^*\tau}.
\end{aligned}
$$

Using the obtained expressions and equation (1.53) we get

$$
\begin{aligned}
G(t,T) \;=\; &\frac{1}{B(t,T)}\exp\left[\bar{x}^*-(\bar{x}^*-x_t)\,e^{-a\tau}+\frac{1}{2}\sigma^2\frac{1-e^{-2a\tau}}{2a}\right]B(t,T) \\
&\cdot\exp\left[-\rho\sigma\theta\left(\frac{1-e^{-a\tau}}{a}-\frac{1-e^{-(a+b)\tau}}{a+b}\right)\right] \\
&\cdot\exp\left[\lambda^*k^*\left(\tau-\frac{1-e^{-a\tau}}{a}\right)\right]. \qquad (1.54)
\end{aligned}
$$

Taking the logarithm of both sides of equation (1.54) yields equation (1.21).

Appendix C: Derivation of formula for the price of a European call option on commodity futures

First, note that equation (1.29) is equivalent to

$$\begin{aligned} C(0,t,T) &= \sum_{n=0}^{\infty} e^{-\lambda^* t} \frac{(\lambda^* t)^n}{n!} \\ &\cdot \mathbb{E}_0^* \left\{ e^{-\int_0^t r_s \mathrm{d}s} \left[F(t,T) - K \;\middle|\; F(t,T) > K,\ \nu_{0,t} = n \right] \right\}. \end{aligned} \tag{1.55}$$

From equations (1.23) and (1.14) we infer that

$$\begin{aligned} &\ln F(t,T) = \ln F(0,T) \\ &+ e^{-a(T-t)} \left[\sigma \int_0^t e^{-a(t-s)} \mathrm{d}W_s^* + \int_0^t e^{-a(t-s)} \ln U_s \mathrm{d}q_s^* \right] - \frac{1}{2}\sigma^2 e^{-2a(T-t)} \frac{1-e^{-2at}}{2a} \\ &+ \lambda^* k^* e^{-aT} \left[\frac{T}{1-e^{-aT}} - \frac{t}{1-e^{-at}} + \frac{e^{-at}}{a(1-e^{-at})} \right]. \end{aligned}$$

Let

$$Q = \ln F(t,T) - \int_0^t r_s \mathrm{d}s.$$

Then

$$\begin{aligned} Q \;=\; \ln F(0,T) \;&+\; e^{-a(T-t)} \left[\sigma \int_0^t e^{-a(t-s)} \mathrm{d}W_s^* - \frac{1}{2}\sigma^2 e^{-a(T-t)} \frac{1-e^{-2at}}{2a} \right] \\ &+\; e^{-a(T-t)} \int_0^t e^{-a(t-s)} \ln U_s \mathrm{d}q_s^* + \lambda^* k^* \left[\frac{(T-t)e^{-a(T-t)}}{1-e^{-a(T-t)}} - \frac{Te^{-aT}}{1-e^{-aT}} \right] \\ &+\; \ln B(0,t) - \frac{\theta^2}{2b^2} \left(t - 2\frac{1-e^{-bt}}{b} + \frac{1-e^{-2bt}}{2b} \right) - \theta \int_0^t \frac{1-e^{-b(t-s)}}{b} \mathrm{d}Z_s^* \end{aligned}$$

so that

$$\begin{aligned} &\mathbb{E}_0^* \left[e^Q \;\middle|\; F(t,T) > K,\ \nu_{0,t} = n \right] \\ &= \mathbb{E}_0^* \left[\exp\left(-\int_0^t r_s \mathrm{d}s \right) \right] \mathbb{E}_0^* \left[e^{\ln F(t,T)} \;\middle|\; F(t,T) > K,\ \nu_{0,t} = n \right] \end{aligned}$$

$$
\begin{aligned}
&\cdot \ \exp\left\{-\mathrm{cov}_0^*\left[\ln F(t,T), \int_0^t r_s \mathrm{d}s \;\middle|\; F(t,T) > K,\ \nu_{0,t} = n\right]\right\} \\
&= B(0,t)\mathbb{E}_0^*\left[e^{\ln F(t,T) + \ln H(t,T)} \;\middle|\; F(t,T) > K,\ \nu_{0,t} = n\right],
\end{aligned}
$$

where

$$
\begin{aligned}
\ln H(t,T) &= -\mathrm{cov}_0^*\left[\ln F(t,T), \int_0^t r_s \mathrm{d}s \;\middle|\; F(t,T) > K,\ \nu_{0,t} = n\right] \\
&= -\rho\sigma\theta e^{-a\tau}\int_0^t e^{-a(t-s)}\mathrm{d}W_s^* \int_0^t \frac{1-e^{-b(t-s)}}{b}\mathrm{d}Z_s^* \\
&= -\rho\sigma\theta\frac{e^{-a\tau}}{b}\left[\frac{1-e^{-at}}{a} - \frac{1-e^{-(a+b)t}}{a+b}\right].
\end{aligned}
$$

Now, $\ln F(t,T) + \ln H(t,T)$ is distributed normally with mean

$$
\ln[F(0,T)H(t,T)] + \varphi(n,t,T) - v(n,\tau)/2
$$

and variance $v(n,\tau)$, where $\varphi(n,t,T)$ and $v(n,\tau)$ are given by equation (1.32) and (1.35), respectively. Hence,

$$
\mathbb{E}_0^*\left[e^Q \;\middle|\; F(t,T) > K,\ \nu_{0,t} = n\right] = B(0,t)F(0,T)H(t,T)e^{\varphi(n,t,T)}N[d_1(n,\tau)], \tag{1.56}
$$

with $d_1(n,\tau)$ being given by equation (1.33).

Similarly, we find that

$$
\begin{aligned}
\mathbb{E}_0^*\left[e^{-\int_0^t r_s \mathrm{d}s}\middle| F(t,T) > K,\ \nu_{0,t} = n\right] &= B(0,t)\mathbb{P}^*\left[F(t,T) > K \;\middle|\; \nu_{0,t} = n\right] \\
&= B(0,t)N[d_2(n,\tau)], \qquad (1.57)
\end{aligned}
$$

where $N[d_2(n,\tau)]$ is defined by equation (1.34). Substituting (1.56) and (1.57) into (1.55) we finally obtain equation (1.30).

Chapter 2

The Role of Jumps in Foreign Exchange Rates

2.1 Introduction

Our knowledge about the complexity of underlying risk factors in exchange rate processes parallels the increase in the number of studies on time series and option prices. The complexity suggests that investment decisions in currency markets will be adequate only if they build upon fairly reasonable specifications of the exchange rate dynamics. Specifically, currency derivatives such as forward rates, options or currency swaps will be very sensitive to volatility dynamics and to higher moments of return distributions.

It is now widely accepted that the exchange rate volatility is time-varying and that the distributions of returns are fat-tailed (see, for example, Bates (1996a,b) and the references cited therein). Figure 2.1, for example, displays the daily relative changes of the exchange rate of Euro with respect to U.S. Dollar, from January 2005 to September 2008. The time-varying nature of volatility is responsible for the interchanging periods of high and low variations in returns. On the other hand, the outliers are manifested through relatively rare but large spikes, or "jumps". The presence of outliers and the extent of skewness are critical for derivatives pricing, as well as hedging and risk management decisions.

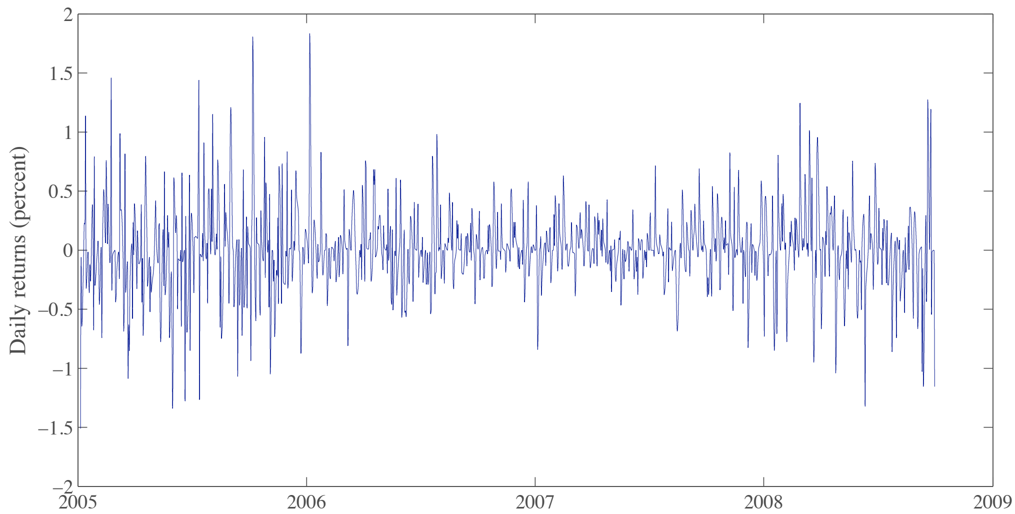

Figure 2.1: **EUR/USD exchange rate.** Daily returns for January 2005–September 2008.

Bates (1988) and Jorion (1988) were among the first to assert that the outliers in exchange rate series can be accounted for by combining a continuous- and a discrete-time process. Many studies have later documented the statistical significance of jumps in exchange rates. Bates (1996b), Jiang (1998), Craine, Lochstoer, and Syrtveit (2000) and Doffou and Hilliard (2001) find that jumps are important components of the currency exchange rate dynamics, even when conditional heteroskedasticity is taken into account. Moreover, several authors had reported that neglecting one of the exchange rate properties usually leads to a significant overestimation of importance of another risk factor (see Jiang (1998) for a discussion). A number of empirical studies revealed other important stylized facts about the exchange rates. For example, Olsen, Müller, Dacorogna, Pictet, Davé, and Guillaume (1997) show that exchange-rate returns in general exhibit non-stable, symmetric, fat-tailed distributions with finite variance and negative first-order autocorrelation and heteroskedasticity.[1]

This chapter studies the nature of jumps in foreign exchange rates, as well as their implications to the option pricing. I propose a general continuous-time stochastic volatility model with Poisson jumps of time-varying intensity. The model conveniently captures all the stylized facts known to the literature. The special cases of the model are several popular benchmarks, such as the Black and Scholes (1973) model, the Merton (1976) model, the stochastic volatility model of Taylor (1986) and the stochastic-volatility jump-diffusion model of Bates (1996b). To estimate the model parameters, I use daily interbank spot exchange rates of Euro, British Pound, Japanese Yen and Swiss Franc with respect to the U.S. Dollar, the four most important exchange rates in terms of currency turnover. The inference framework is based on the efficient method of moments procedure of Gallant and Tauchen (1996).

The results confirm that both stochastic volatility and jumps play a critical role in the exchange rate dynamics. Moreover, a correctly specified model should include a bimodal distribution of jump sizes. Depending on the exchange rate, a model with the volatility-dependent jump intensity may outperform a model with a constant intensity. The proposed general model also allows for a closed-form solution for the price of European-style currency options. It is capable to accommodate the shapes of Black-Scholes implied volatilities observed in the actual data. This indicates that the dominant empirical characteristics of exchange rate processes seem to be priced by the market.

The remainder of the chapter is organized as follows: Section 2.2 develops a model specification for exchange rates and describes the estimation methodology. Section 2.3 describes the data and provides the estimation results. Section 2.4 considers the option pricing implications of jumps. Concluding remarks

[1] The economic literature dealing with jump processes and their pricing implications has been growing ever since the seminal work of Merton (1976). Examples include Ball and Torous (1985), Bates (1991), Bates (1996a), Bates (1996b), Chernov, Gallant, Ghysels, and Tauchen (1999), Pan (2002) and Andersen, Benzoni, and Lund (2002).

are given in Section 2.5.

2.2 Model Specification and Estimation Methodology

2.2.1 Model

The model is constructed to capture the salient features of exchange rate dynamics and incorporate the majority of popular models used in the literature as its special cases. I will assume that the instantaneous exchange rate S_t solves

$$\frac{\mathrm{d}S_t}{S_t} = \mu \mathrm{d}t + \sqrt{V_t}\,\mathrm{d}W_{1,t} + (e^{u_t} - 1)\,\mathrm{d}q_t - \lambda_t \bar{k}\mathrm{d}t, \tag{2.1}$$

where the instantaneous variance V_t follows a mean-reverting diffusion given by the "square-root" specification of Heston (1993):

$$\mathrm{d}V_t = (\alpha - \beta V_t)\,\mathrm{d}t + \sigma\sqrt{V_t}\,\mathrm{d}W_{2,t}. \tag{2.2}$$

The stochastic processes $W_{1,t}$ and $W_{2,t}$ are standard Brownian motions on the usual probability-space triple $(\Omega, \mathcal{F}_t, \mathbb{P})$, where $\mathbb{P}$ is the "physical", or the data-generating measure. The correlation between $W_{1,t}$ and $W_{2,t}$ is ρ, which can be written as

$$\mathrm{d}W_{1,t}\mathrm{d}W_{2,t} = \rho\mathrm{d}t. \tag{2.3}$$

The term $(e^{u_t} - 1)\,\mathrm{d}q_t$ in equation (2.1) is the jump component. The returns jump at t if the Poisson counter (or jump "flag") $\mathrm{d}q_t$ is equal to one, which happens with probability $\lambda_t\mathrm{d}t$. Jump intensity λ_t may change over time. In particular, jumps may be more likely in periods of high volatility. I will therefore allow the intensity to be a linear function of the instantaneous variance,

$$\lambda_t = \lambda_0 + \lambda_1 V_t. \tag{2.4}$$

The random variable u_t in equation (2.1) determines the relative magnitude of a jump. The processes $\mathrm{d}q_t$ and u_t are independent, both are serially uncorrelated, and both are uncorrelated with diffusions $\mathrm{d}W_{1,t}$ and $\mathrm{d}W_{2,t}$. Also, neither $\mathrm{d}q_t$ nor u_t are measurable with respect to $\mathcal{F}_t$.

It is reasonable to assume that distribution of jump sizes is *not* concentrated around zero. This is actually not the case in most of the jump-diffusion specifications in the literature: jump sizes are usually modeled as random variables from a unimodal distribution. Since jumps can be both positive and negative, their unconditional expected size is typically close to zero. Unimodal jump-size distributions imply that majority of jumps will be relatively small in magnitude, which is exactly the opposite of their nature. They will also tend to increase kurtosis by adding more mass at the center of the return distribution instead of adding it to the tails. In this way, the effect of fat tails is achieved through normalization of the probability density function. In such specifications, most of the jumps are difficult to distinguish from returns generated by diffusion, which may lead to an overestimation of jump frequencies. Johannes (2004), for example, estimates a jump-diffusion interest rate model and finds jump intensities that are between 0.05 and 0.10, but detects only 5 jumps per year, which corresponds to an intensity of around 0.02.

I will therefore assume that the variable u_t, which determines the size of the jump, comes from a mixture of two normal distributions, one centered around a positive value, the other around a negative value:

$$u_t \;\sim\; p\,\mathcal{N}\left(\ln(1+k)-\omega^2/2,\;\omega^2\right) \;+\; (1-p)\,\mathcal{N}\left(\ln(1-k)-\omega^2/2,\;\omega^2\right). \quad (2.5)$$

Hence, p has the meaning of the probability that the jump is positive, k is the expected size of a positive jump, while $-k$ is the expected size of a negative jump. At time t, the expected contribution of jumps to return $\mathrm{d}S_t/S_t$ is

$$\mathbb{E}_t\left[\left(e^{u_t}-1\right)\mathrm{d}q_t\right] = \lambda_t\bar{k}\mathrm{d}t,$$

where

$$1 + \bar{k} \equiv p(1 + k) + (1 - p)(1 - k).$$

Therefore, the return process is constructed such that the jumps are on average compensated by the last term in equation (2.1). I use $\mathbb{E}_t(\cdot)$ to denote the conditional expectation given the information available at time t, instead of a more cumbersome $\mathbb{E}(\cdot|\mathcal{F}_t)$.

The outlined model specification has a form of a stochastic volatility jump-diffusion process with bimodal distribution of jump sizes (hereafter: SVJD-B).[2] It has a convenient feature that it contains several popular jump- and pure-diffusion benchmark models as its special cases. For example, by setting $p = 1$ and $\lambda_1 = 0$ we obtain the usual SVJD specification of Bates (1996b) or Bates (2000). A stochastic volatility model without jumps (SV) of Taylor (1986) is obtained by setting all jump parameters (λ_0, λ_1, p, k and ω) to zero. Merton (1976) diffusion model with constant variance is obtained by setting all stochastic-volatility parameters (α, β, ρ, λ_1) to zero, introducing a constant jump intensity ($\lambda_t = \lambda_0$, $\lambda_1 = 0$) and constraining the distribution of jump sizes to be unimodal ($p = 1$). Finally, the Black and Scholes (1973) model (BS) is obtained by setting all jump parameters to zero, α, β and ρ to zero, and (with a slight abuse of notation) by fixing $V_t = \sigma^2$.

2.2.2 Estimation Methodology

Estimation of a continuous-time model, such as one given by equations (2.1)–(2.2), is never straightforward when we bring it to discretely sampled data. The main difficulty lies in the fact that closed-form expressions for a discrete transition density are seldom available. The presence of unobservable state variables, such as stochastic volatility, makes this task even more arduous. In principle, some form of maximum likelihood estimation might be feasible (see,

[2] To the best of my knowledge, the bimodal assumption for the distribution of jump sizes was previously used only in a numerical valuation of real options in Dias and Rocha (2001).

for example, Lo (1988)), but it is based on computationally very demanding numerical procedures that involve integration of latent variables out of the likelihood function. The problem becomes even more difficult when jumps are introduced into the model.

A number of alternatives to the maximum likelihood technique have been proposed to overcome the issue of computational inefficiency. Examples of simulation-based inference for jump-diffusion models can be found in Andersen, Benzoni, and Lund (2002), Duffie, Pan, and Singleton (2000) and Chernov, Gallant, Ghysels, and Tauchen (1999). Simulation approaches based on the method of moments are a useful tool whenever it is possible to alleviate the problem of inefficient inference, which can be done by careful selection of moment conditions. For example, Pan (2002) uses the simulated method of moments (SMM) of Duffie and Singleton (1993) and matches sample moments with the simulated ones to estimate risk premia embedded in options on a stock market index. The efficient method of moments (EMM) of Gallant and Tauchen (1996) refines the SMM approach by a convenient choice of moment conditions: they are obtained from the expected score of the auxiliary model. The auxiliary model is a discrete-time model whose purpose is to approximate the sample distribution. Hence, there are at least two good features of the EMM approach: first, it will achieve the efficiency of the maximum likelihood technique under reasonable assumptions, and second, the objective function can be used to test for overidentifying restrictions, as with an ordinary generalized method of moments.

Several jump-diffusion models were developed to describe the exchange rate dynamics. Bates (1996b), for example, estimates the parameters of an SVJD model from the prices of Deutsche Mark options traded on the Philadelphia Stock Exchange. More recently, Maheu and McCurdy (2006) proposed a discrete-time model of foreign exchange rate returns with jumps. Their estimation is based on a Markov Chain Monte Carlo technique. Although this method is a powerful inference tool, its implementation always has to be tailored for a particular choice of model, making it difficult to compare with other specifications.

I use the EMM to estimate the proposed SVJD-B model (2.1)–(2.2) and to compare it with the alternatives. As pointed out by Andersen, Benzoni, and Lund (2002), the EMM procedure critically relies on the correct specification of the auxiliary model. The auxiliary model should approximate the conditional distribution of the return process as close as possible. If the score of the auxiliary model asymptotically spans the score of the true model, the EMM will be asymptotically efficient (see Gallant and Long (1997) for the proof). Therefore, any auxiliary model should capture the dominant features of the return dynamics in a discrete-time series. Specifically, it should be able to take into account the presence of autocorrelation and heteroskedasticity, as well as to model any excess skewness and kurtosis. A semi-nonparametric (SNP) specification for the auxiliary model by Gallant and Nychka (1987) is based on the notion that higher-order moments of distribution can be captured with a polynomial expansion.

Given that a set of data is stationary, an ARMA term is sufficient to describe the conditional mean, while an ARCH-type term should be able to filter out conditional heteroskedasticity. I choose the EGARCH model of Nelson (1991) in order to capture both heteroskedasticity and potential presence of asymmetric responses of conditional variance to positive and negative returns. Finally, to accommodate the presence of fat tails in the return distribution, I augment the conditional probability density function of the auxiliary model by a polynomial in standardized returns.

The semi-nonparametric (SNP) estimation step is performed via quasi-maximum likelihood technique on the fully specified auxiliary model. I follow Andersen, Benzoni, and Lund (2002) and assume that auxiliary model follows an ARMA(r,m)-EGARCH(p,q)-Kz(K_z)-Kx(K_x) process with a probability distribution function of the form:

$$f_K(y_t|\mathcal{F}_{t-1};\varphi) = \frac{[P_K(z_t,x_t)]^2}{\int[P_K(z,x)]^2\phi(z)\mathrm{d}z}\frac{\phi(z_t)}{\sqrt{h_t}} \tag{2.6}$$

where $y_t \equiv \ln(S_t/S_{t-1})$ is a vector of log-returns that follows an ARMA(r,m) process

$$y_t = \mu + \sum_{i=1}^{r} b_i y_{t-i} + \varepsilon_t + \sum_{i=1}^{m} c_i \varepsilon_{t-i}. \tag{2.7}$$

The residuals ε_t are assumed to be normally distributed conditionally on the information available one time step before:

$$\varepsilon_t | \mathcal{F}_{t-1} \sim \mathcal{N}(0, h_t). \tag{2.8}$$

The corresponding standardized residuals are $z_t = \varepsilon_t/\sqrt{h_t}$, and x_t is the vector of their lags. The standard normal probability density function is labeled by $\phi(\cdot)$. The conditional variance h_t follows an EGARCH(p,q) process of the form

$$\ln h_t = \omega + \sum_{i=1}^{p} \beta_i \ln h_{t-i} + \sum_{j=1}^{q} \alpha_j \left(|z_{t-j}| - \sqrt{\frac{2}{\pi}} \right) + \sum_{j=1}^{q} \theta_j z_{t-j}. \tag{2.9}$$

In equation (2.6), the full set of parameters is labeled by φ. Finally, $P_K(\cdot)$ is a nonparametric polynomial expansion given by

$$P_K(z, x) = \sum_{i=0}^{K_z} \sum_{j=0}^{K_x} \left(a_{ij} x^j \right) z^i, \quad a_{00} = 1. \tag{2.10}$$

Here, as in Andersen, Benzoni, and Lund (2002), the coefficients in expansion depend on lags x. This expansion is designed to capture any excess kurtosis in returns, but also to accommodate additional skewness that has not already been represented by the EGARCH term. I use the Bayesian information criterion (BIC) to select the best fitting model for each series.

The EMM estimation step works in the following way. Given the set of parameters

$$\psi = \{\mu, \alpha, \beta, \sigma, \rho, \lambda_0, \lambda_1, p, k, \omega\},$$

I simulate the sample of exchange rates $\{\widetilde{S}_t\}_{t=1}^{T_{\text{sim}}}$ and instantaneous variances $\{\widetilde{V}_t\}_{t=1}^{T_{\text{sim}}}$ using the specification given by the continuous-time model (2.1)–(2.2).

The EMM estimator of model parameters ψ is defined as

$$\widehat{\psi} = \arg\min_{\psi} \; \mathbf{m}(\psi, \widehat{\varphi})' \; \mathbf{W} \; \mathbf{m}(\psi, \widehat{\varphi}), \tag{2.11}$$

where $\mathbf{m}(\psi, \widehat{\varphi})$ is the expectation of the score function and $\widehat{\varphi}$ is the quasi-maximum likelihood estimate of the set of SNP parameters. The expectation of the score is evaluated as the sample mean across simulations,

$$\mathbf{m}(\psi, \widehat{\varphi}) = \frac{1}{T_{\text{sim}}} \sum_{t=1}^{T_{\text{sim}}} \frac{\partial \ln f_K(\widetilde{y}_t | \mathcal{F}_{t-1}; \widehat{\varphi})}{\partial \varphi},$$

where $\widetilde{y}_t \equiv \ln(\widetilde{S}_t / \widetilde{S}_{t-1})$. The weighting matrix $\mathbf{W}$ is a consistent estimate of the inverse asymptotic covariance matrix of the auxiliary score.

To reduce the effects of discretization, I sample at time intervals of 1/10 of a day. At each run, two antithetic samples were created for the purpose of variance reduction, each of length $100,000 \times 10 + 20,000$. To eliminate the effects of initial conditions, I discard the "burn-in" period of the first 20,000 simulated points. The final sample of $T_{\text{sim}} = 100,000$ daily log-returns, $\{\widetilde{y}_t\}_{t=1}^{T_{\text{sim}}}$, was obtained by adding up the groups of 10 elements in the simulated sample.

2.3 Estimation Results

2.3.1 Data

The results are based on average daily interbank spot exchange rates of Euro, British Pound, Japanese Yen and Swiss Franc with respect to the U.S. Dollar, from January 4, 1999 to September 30, 2008, a sample of 2542 observations. All four time series, obtained from Thomson Financial's Datastream, are shown in Figure 2.2. The JPY/USD exchange rate is expressed per 100 Yens. Table 3.1 provides summary statistics for the exchange rate levels S_t and the corresponding daily returns, computed as $y_t = \ln(S_t/S_{t-1})$. Daily sampling is chosen in order to capture high-frequency fluctuations in return processes that may be critical for identification of jump components, while avoiding to model the intraday return dynamics, abundant with spurious market microstructure distortions and trading frictions.

Table 2.1: **Summary Statistics**
Daily interbank spot exchange rates of Euro, British Pound, Japanese Yen and Swiss Franc with respect to the U.S. Dollar, from January 4, 1999 to September 30, 2008 (2542 observations).

Panel A: Daily exchange rate levels				
Currency	Mean	Variance	Skewness	Kurtosis
EUR	1.1511	0.0376	0.2234	2.1992
GBP	1.7103	0.0376	0.0661	1.7385
JPY	0.8774	0.0030	−0.0339	2.2511
CHF	0.7380	0.0121	0.1375	2.1955

Panel B: Daily returns (percent)				
Currency	Mean	Variance	Skewness	Kurtosis
EUR	0.0084	0.3539	−0.0267	4.5420
GBP	0.0040	0.2338	0.0757	4.1778
JPY	0.0026	0.3493	0.2267	4.8656
CHF	0.0088	0.4012	0.1411	4.2532

I perform several preliminary test on the data. The values of skewness and kurtosis in Table 3.1 indicate that both the levels and returns deviate from normality. This is also confirmed by Jarque-Bera and Kolmogorov-Smirnov tests (not reported), whose p-values are at most of the order of 10^{-3}. Table 3.2 shows the results of Ljung-Box test for the autocorrelation of returns, up to order 10 (Panel A). The null hypotheses of no autocorrelation in returns cannot be rejected. The absence of a significant short-run return predictability is consistent with high efficiency of the currency market. The autocorrelation in the squared returns is, on the other hand, highly significant in all four series, indicating the presence of heteroskedasticity (Panel B). The correlation coef-

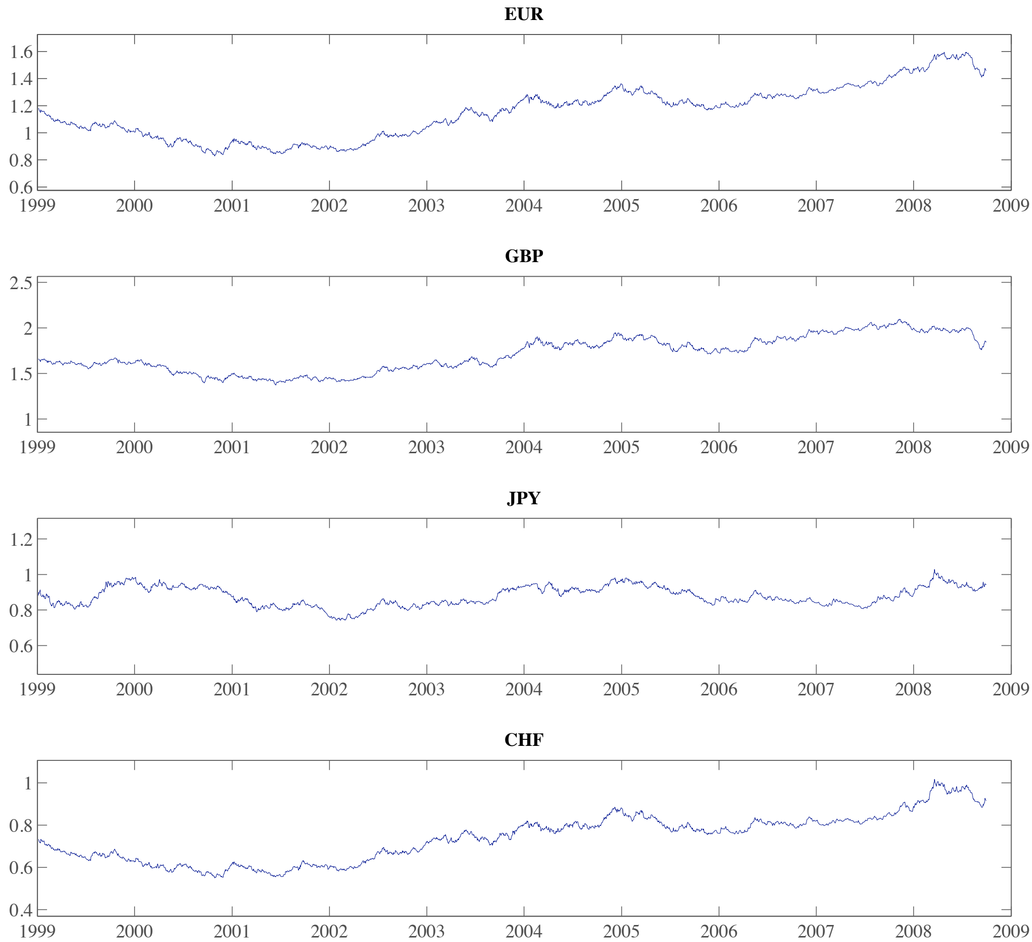

Figure 2.2: **Daily exchange rate levels:** January 4, 1999 to September 30, 2008. The JPY/USD rate is expressed per 100 Yens.

ficients between squared returns and their lags (not reported) are all positive, confirming the notion of clustering – the periods of high volatility are likely to be followed by high volatility.

Table 3.3 reports the results of the unit root tests. The values of the Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) statistics indicate that the unit root hypothesis is convincingly rejected in favor of stationary returns (the critical values of ADF and PP statistics at 5 and 1 percent confidence

Table 2.2: **Autocorrelation**
Ljung-Box test for autocorrelation of returns and squared returns up to 10^{th} lag.

Panel A: Autocorrelation of returns		
Currency	Q statistic	p-value
EUR	3.9867	0.9479
GBP	9.4858	0.4867
JPY	6.8611	0.7385
CHF	12.7326	0.2390

Panel B: Autocorrelation of squared returns		
Currency	Q statistic	p-value
EUR	111.5435	$< 10^{-5}$
GBP	105.7946	$< 10^{-5}$
JPY	81.5108	$< 10^{-5}$
CHF	42.7107	$< 10^{-5}$

are −3.41 and −3.96, respectively). The stationarity is a prerequisite for any method of moments approach that is critically relying on stability of the data-generating process.

Finally, I report the results of the Jiang and Oomen (2008) swap-variance test for detection of jumps in returns and squared returns, Table 2.4. The swap-variance test exploits the impact of jumps on the third and higher order moments of asset returns. The test is based on the statistic

$$\frac{T}{\sqrt{\Omega}}\left(SwV_T - RV_T\right) \;\sim\; \mathcal{N}(0,1),$$

where

$$SwV_T = 2\sum_{t=2}^{T}\left(R_t - y_t\right)$$

is twice the accumulated difference between discretely and continuously com-

Table 2.3: **Stationarity**
Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests for the presence of unit roots, based on the regression

$$y_t = c + \delta t + \phi y_{t-1} + \sum_{L=1}^{10} b_L \Delta y_{t-L} + \varepsilon_t,$$
$$\mathrm{H}_0 : \phi = 1,\ \delta = 0.$$

Currency	ADF statistic	PP statistic
EUR	−15.8257	−50.4768
GBP	−15.2648	−48.0508
JPY	−14.6802	−49.2301
CHF	−15.6345	−50.9751
5% crit. value	−3.41	−3.41
1% crit. value	−3.96	−3.96

pounded returns, $R_t = S_t/S_{t-1} - 1$ and $y_t = \ln(S_t/S_{t-1})$, respectively,

$$RV_T = \sum_{t=2}^{T} y_t^2$$

is the realized variance, and Ω is the asymptotic variance of the test statistic. The robust estimator of Ω is given by

$$\widehat{\Omega} = \frac{1}{9}\frac{\mu_6}{\mu_1^6}\frac{T^3}{T-5}\sum_{s=1}^{T-6}\prod_{t=1}^{6}|y_{s+t}|,$$

with $\mu_k = \mathbb{E}(|y|^k)$. The null hypothesis of the swap-variance test is that S_t follows a process without jumps. Intuitively, the test statistic reflects the cumulative gain of a variance swap replication strategy which is known to be minimal in the absence of jumps but substantial in the presence of jumps. If the underlying process is continuous, the difference between SwV_T and RV_T

should asymptotically go to zero. The results of the test indicate that in all four series the jumps in returns are highly significant (Panel A), whereas the jumps in squared returns are not (Panel B). This implies that it is not necessary to overparameterize the model by introducing discontinuities into the process for conditional variance.

Table 2.4: **Presence of jumps**
Swap-variance jump test, based on the test statistics

$$\frac{T}{\sqrt{\widehat{\Omega}}}\left(SwV_T - RV_T\right) \;\sim\; \mathcal{N}(0,1),$$

where

$$SwV_T = 2\sum_{t=2}^{T}\left(R_t - y_t\right), \quad RV_T = \sum_{t=2}^{T} y_t^2,$$
$$R_t = S_t/S_{t-1} - 1, \quad y_t = \ln(S_t/S_{t-1}),$$
$$\widehat{\Omega} = \frac{1}{9}\frac{\mu_6}{\mu_1^6}\frac{T^3}{T-5}\sum_{s=1}^{T-6}\prod_{t=1}^{6}|y_{s+t}|, \quad \mu_k = \mathbb{E}(|y|^k),$$

H_0 : S_t follows a process without jumps.

Panel A: Jumps in returns

Currency	Swap-var stat.	p-value
EUR	-6.9228	$< 10^{-5}$
GBP	-10.9712	$< 10^{-5}$
JPY	-5.9042	$< 10^{-5}$
CHF	-7.9779	$< 10^{-5}$

Panel B: Jumps in squared returns

Currency	Swap-var stat.	p-value
EUR	-0.0254	0.9797
GBP	-0.0259	0.9794
JPY	-0.0220	0.9824
CHF	-0.0210	0.9833

2.3.2 Estimation of the auxiliary model

Table 2.5 reports the results of the SNP step. It shows the quasi-maximum likelihood parameter estimates $\widehat{\varphi}$, along with their standard errors. I ran the estimations across the possible combinations (r, m, p, q, K_x, K_z), allowing each of the parameters to take values between 0 and 10. The selection criterion based on the BIC indicates that the best-fitting auxiliary models have the form ARMA(0,0)-EGARCH(1,1)-Kz(K_z)-Kx(0), with K_z being 8, 7, 6 and 7 for the Euro, Pound, Yen and Franc exchange rate, respectively.[3] Table 2.5 also shows the total number of SNP parameters n, as well as the optimal values of log-likelihood functions, LL. The absence of the ARMA term is not surprising given that the data exhibit no significant autocorrelation. Also, in all four cases heteroskedasticity is entirely captured by the first lags of conditional variance and return innovations in the EGARCH terms, as $p = 1$ and $q = 1$. The values of EGARCH parameter β governing the persistence are close to the boundary of covariance stationary region, but still significantly within the boundaries. The parameter θ is relatively small and – with the exception of Swiss Franc – statistically insignificant. This indicates that the "leverage" effect does not play such an important role in dynamics of exchange rates. The terms Kx that should take into account heterogeneity in the polynomial expansion are always insignificant ($K_x = 0$), which indicates that the EGARCH leading terms pick up all the serial dependence in the returns.

[3]The actual values of BIC are not reported, but are available upon request.

Table 2.5: **Estimates of the Auxiliary Model**

	EUR	GBP	JPY	CHF
μ	0.0558	0.0844	−0.0588	0.0737
	(0.0350)	(0.0493)	(0.0443)	(0.0434)
ω	−0.0065	−0.0437	−0.0263	−0.0021
	(0.0034)	(0.0162)	(0.0242)	(0.0035)
β	0.9945	0.9799	0.9731	0.9969
	(0.0018)	(0.0078)	(0.0093)	(0.0016)
α	0.0675	0.0583	0.1239	0.0401
	(0.0119)	(0.0149)	(0.0288)	(0.0096)
θ	0.0047	−0.0023	0.0203	0.0161
	(0.0063)	(0.0080)	(0.0144)	(0.0064)
a_{10}	−0.1275	−0.1671	0.0684	−0.1815
	(0.0588)	(0.0776)	(0.0646)	(0.0606)
a_{20}	−0.1274	−0.0274	−0.1407	−0.0980
	(0.0521)	(0.0582)	(0.1172)	(0.0632)
a_{30}	0.0743	0.0701	−0.0187	0.0852
	(0.0240)	(0.0206)	(0.0161)	(0.0206)
a_{40}	0.0330	0.0119	0.0237	0.0205
	(0.0184)	(0.0122)	(0.0237)	(0.0130)
a_{50}	−0.0141	−0.0130	0.0024	−0.0126
	(0.0039)	(0.0033)	(0.0014)	(0.0035)
a_{60}	−0.0013	0.0004	−0.0004	−0.0002
	(0.0024)	(0.0009)	(0.0017)	(0.0009)
a_{70}	0.0008	0.0006		0.0005
	(0.0002)	(0.0002)		(0.0002)
a_{80}	0.0000			
	(0.0001)			
Model	001180	001170	001160	001170
n	13	12	11	12
LL	9597.85	10048.37	9568.58	9402.97

(Standard errors in parentheses.)

2.3.3 EMM estimation

Once we have the optimal score parameters $\widehat{\varphi}$ obtained from the SNP model, we can estimate the main step of the EMM procedure. The EMM parameter estimates $\widehat{\psi}$ obtained from the SVJD-B and the competing models are summarized in Tables 2.6–2.9. Standard errors are given in the parentheses. Tables also report the results of Hansen's test of overidentifying restrictions: chi-squares, degrees of freedom and p-values.

We can draw several important conclusions from these estimates. First, as expected, stochastic volatility is important: the constant-volatility Merton and Black-Scholes models can be overwhelmingly rejected in all four cases. Second, jumps are statistically significant, since SVJD-B and SVJD specifications outperform the SV model without jumps. The SV model is also strongly rejected at any reasonable level of significance for Euro and Franc. Third, the usual SVJD specification is outperformed by the SVJD-B model with bimodal distribution of jump sizes given by equation (2.5). The SVJD model may as well be rejected at significance levels less than 0.05. Fourth, the dependence of jump intensity on volatility levels as given by the affine specification (2.4) is important, but the alternative of constant jump intensity ($\lambda_1 = 0$) cannot be easily rejected. For example, for the Yen exchange rate the restricted model is significant at 0.05 level, while the fully specified SVJD-B model is not. The constant term λ_0 is by an order of magnitude greater than the affine coefficient λ_1. Finally, the correlation between return and volatility is important: the estimated values of ρ are significant and negative. The restriction $\rho = 0$ may not be rejected only for the Euro exchange rate. As suggested by Andersen, Benzoni, and Lund (2002), a negative correlation between return diffusion and volatility can explain part of the skewness in returns.

Table 2.6: **EMM Estimates: EUR**

Parameter	SVJD-B	SVJD-B $\lambda_1 = 0$	SVJD-B $\rho = 0$	SVJD	SV	Merton	BS
μ	0.0239	0.0287	0.0270	0.0516	0.0573	0.0558	0.0558
	(0.0097)	(0.0029)	(0.0044)	(0.0022)	(0.0025)	(0.0031)	(0.0043)
α	0.0021	0.0021	0.0002	0.0002	0.0019		
	(0.0033)	(0.0061)	(0.0136)	(0.0069)	(0.0019)		
β	0.0063	0.0055	0.0006	0.0018	0.0055		
	(0.0035)	(0.0025)	(0.0041)	(0.0031)	(0.0078)		
σ	0.1078	0.1070	0.0683	0.8174	0.0992	0.4418	0.4453
	(0.0063)	(0.0026)	(0.0049)	(0.0026)	(0.0033)	(0.0044)	(0.0050)
ρ	−0.0255	−0.0262		−0.0236	−0.0262		
	(0.0044)	(0.0044)		(0.0044)	(0.0021)		
λ_0	0.0298	0.0308	0.0248	0.0247		0.0308	
	(0.0053)	(0.0037)	(0.0038)	(0.0060)		(0.0043)	
λ_1	0.0024		0.0000				
	(0.0066)		(0.0136)				
p	0.5227	0.4813	0.3844				
	(0.0153)	(0.0165)	(0.0025)				
k	1.4834	1.4958	1.1948	1.2596		−0.0582	
	(0.0142)	(0.0038)	(0.0045)	(0.0062)		(0.0073)	
ω	0.4413	0.4876	0.3771	0.5271		0.4714	
	(0.0030)	(0.0043)	(0.0022)	(0.0029)		(0.0035)	
χ^2 [$d.f.$]	7.5852 [3]	10.5561 [4]	9.0277 [4]	13.3386 [5]	23.3928 [8]	22.5033 [8]	96.5893 [11]
(p-value)	(0.0554)	(0.0320)	(0.0604)	(0.0204)	(0.0029)	(0.0041)	$< 10^{-5}$

(Standard errors in parentheses.)

Table 2.7: **EMM Estimates: GBP**

Parameter	SVJD-B	SVJD-B $\lambda_1 = 0$	SVJD-B $\rho = 0$	SVJD	SV	Merton	BS
μ	0.0198	0.0174	0.0277	0.0180	0.1206	0.0844	0.0844
	(0.0138)	(0.0049)	(0.0035)	(0.0123)	(0.0021)	(0.0021)	(0.0026)
α	0.0011	0.0026	0.0019	0.0039	0.0034		
	(0.0099)	(0.0039)	(0.0065)	(0.0108)	(0.0029)		
β	0.0139	0.0263	0.0182	0.0534	0.0448		
	(0.0187)	(0.0061)	(0.0075)	(0.0082)	(0.0025)		
σ	0.2099	0.4469	0.3699	0.7219	0.2027	0.2167	0.2179
	(0.0195)	(0.0067)	(0.0028)	(0.0047)	(0.0026)	(0.0032)	(0.0043)
ρ	-0.0140	-0.0157		-0.0151	-0.0157		
	(0.0205)	(0.0082)		(0.0038)	(0.0016)		
λ_0	0.0300	0.0291	0.0280	0.0312		0.0324	
	(0.0209)	(0.0082)	(0.0082)	(0.0023)		(0.0030)	
λ_1	0.0037		0.0038				
	(0.0091)		(0.0046)				
p	0.3746	0.4760	0.4513				
	(0.0202)	(0.0205)	(0.0030)				
k	1.4241	1.1564	1.1884	1.0728		-0.1800	
	(0.0057)	(0.0060)	(0.0045)	(0.0063)		(0.0124)	
ω	0.3337	0.3516	0.2769	0.3147		0.3192	
	(0.0036)	(0.0034)	(0.0032)	(0.0036)		(0.0038)	
χ^2 [$d.f.$]	4.4774 [2]	7.1153 [3]	14.1125 [3]	10.3601 [4]	17.1286 [7]	32.0357 [7]	28.1604 [10]
(p-value)	(0.1066)	(0.0683)	(0.0028)	(0.0348)	(0.0166)	$< 10^{-5}$	(0.0017)

(Standard errors in parentheses.)

Table 2.8: **EMM Estimates: JPY**

Parameter	SVJ-B	SVJD-B $\lambda_1 = 0$	SVJD-B $\rho = 0$	SVJD	SV	Merton	BS
μ	0.0128	-0.0048	0.0459	-0.0700	-0.0305	-0.0588	-0.0589
	(0.0215)	(0.0104)	(0.0900)	(0.0500)	(0.0240)	(0.0250)	(0.0280)
α	0.0112	0.0044	0.0126	0.0127	0.0030		
	(0.0181)	(0.0088)	(0.0025)	(0.0047)	(0.0045)		
β	0.0880	0.0446	0.0466	0.0931	0.0146		
	(0.0182)	(0.0081)	(0.0038)	(0.0036)	(0.0030)		
σ	0.3473	0.1729	0.2114	0.3656	0.0779	0.3064	0.3068
	(0.0106)	(0.0075)	(0.0062)	(0.0034)	(0.0024)	(0.0026)	(0.0028)
ρ	-0.0276	-0.0287		-0.0278	-0.0263		
	(0.0115)	(0.0082)		(0.0037)	(0.0033)		
λ_0	0.0263	0.0242	0.0262	0.0261		0.0260	
	(0.0059)	(0.0078)	(0.0085)	(0.0036)		(0.0090)	
λ_1	0.0045		0.0021				
	(0.0208)		(0.0081)				
p	0.6754	0.6616	0.5810				
	(0.0488)	(0.0519)	(0.0121)				
k	1.5733	1.6358	1.8133	1.7183		0.3912	
	(0.0270)	(0.0044)	(0.0057)	(0.0065)		(0.0061)	
ω	0.5099	0.4851	0.4355	0.4008		0.4535	
	(0.0102)	(0.0262)	(0.0128)	(0.0143)		(0.0134)	
χ^2 [$d.f.$]	4.4367 [1]	5.3276 [2]	9.4297 [2]	9.2045 [3]	14.3776 [6]	17.6237 [6]	24.3500 [9]
(p-value)	(0.0352)	(0.0697)	(0.0090)	(0.0267)	(0.0257)	(0.0072)	(0.0038)

(Standard errors in parentheses.)

Table 2.9: **EMM Estimates: CHF**

Parameter	SVJD-B	SVJD-B $\lambda_1 = 0$	SVJD-B $\rho = 0$	SVJD	SV	Merton	BS
μ	0.0077	0.0190	0.0424	0.0056	0.0259	0.0737	0.0743
	(0.0354)	(0.0038)	(0.0054)	(0.0051)	(0.0019)	(0.0020)	(0.0022)
α	0.0042	0.0012	0.0009	0.0007	0.0011		
	(0.0108)	(0.0055)	(0.0051)	(0.0028)	(0.0042)		
β	0.0147	0.0100	0.0115	0.0017	0.0027		
	(0.0128)	(0.0088)	(0.0035)	(0.0088)	(0.0046)		
σ	0.6100	0.4101	0.0851	0.1175	0.0935	0.1118	0.1128
	(0.0193)	(0.0060)	(0.0028)	(0.0037)	(0.0036)	(0.0039)	(0.0041)
ρ	−0.0216	−0.0215		−0.0210	−0.0220		
	(0.0093)	(0.0048)		(0.0027)	(0.0026)		
λ_0	0.0356	0.0361	0.0324	0.0310		0.0344	
	(0.0204)	(0.0040)	(0.0042)	(0.0026)		(0.0043)	
λ_1	0.0059		0.0041				
	(0.0365)		(0.0022)				
p	0.5625	0.6271	0.5249				
	(0.0592)	(0.0641)	(0.0038)				
k	1.5652	1.5771	1.6466	1.5651		0.1941	
	(0.0050)	(0.0049)	(0.0070)	(0.0074)		(0.0074)	
ω	0.4053	0.4192	0.4230	0.3967		0.4408	
	(0.0355)	(0.0058)	(0.0057)	(0.0062)		(0.0061)	
χ^2 [$d.f.$]	5.1439 [2]	9.4681 [3]	15.2722 [3]	10.3886 [4]	25.5051 [7]	20.8516 [7]	47.2409 [10]
(p-value)	(0.0764)	(0.0237)	(0.0016)	(0.0344)	(0.0006)	(0.0040)	$< 10^{-5}$

(Standard errors in parentheses.)

The values of the leading intensity term λ_0 roughly indicate that jumps should on average occur between 7 and 10 times per year, depending on the exchange rate. Although jumps are rare, their significance implies that they cannot be ruled out. Positive jumps are more likely on average, with the exception of the British Pound, where about 63 percent of jumps are negative. The unconditional mean of jump sizes, $\bar{k} = (2p-1)k$, is close to zero and positive, except for the Pound, where $p < 0.5$. This asymmetry captures a part of the skewness of the unconditional return distribution. The confidence bounds for jump sizes can be obtained from the values of the standard deviation ω. For example, positive jumps in the Euro exchange rate happen with probability 0.52 and have magnitudes that are in the 95-percent confidence interval of [0.61, 2.36] percent.

Using the estimated parameters, we can infer the ex-post probability of a jump on a given date implied by the actual data. Following Johannes (2004), I use a Gibbs sampling technique to compute the filtering distribution of jump times and jump sizes. The Gibbs sampler iteratively samples from the filtering distribution of variances

$$\pi(V_{t+\Delta t}|V_t, q_{t+\Delta t}, u_{t+\Delta t}, y_{t+\Delta t}, y_t; \widehat{\psi}),$$

the filtering distribution of jump times

$$\pi(q_{t+\Delta t}|u_{t+\Delta t}, y_{t+\Delta t}, y_t, V_{t+\Delta t}, V_t; \widehat{\psi}),$$

and the filtering distribution of jump sizes

$$\pi(u_{t+\Delta t}|q_{t+\Delta t}, y_{t+\Delta t}, y_t, V_{t+\Delta t}, V_t; \widehat{\psi}),$$

all of which are know distributions, where $\{y_t\}_{t=1}^T$ is the observed time series of daily returns and $\widehat{\psi}$ are the estimated SVJD-B parameters. In each iteration j, the algorithm produces a sequence $\{\{V_t^{(j)}\}_{t=1}^T,\ \{q_t^{(j)}\}_{t=1}^T,\ \{u_t^{(j)}\}_{t=1}^T\}$ of conditional variances, jump flags and jump sizes, which are draws from the joint distribution $\pi(V_{t+\Delta t}, q_{t+\Delta t}, u_{t+\Delta t}|y_{t+\Delta t}, y_t, V_t; \widehat{\psi})$. The algorithm converges quickly since there is no parameter uncertainty. Hence, I work with at most $10,000$ iteration steps and discard the "burn-in" period of the first 2,000 iterations.

Figures 2.3–2.6 display the results. They show daily returns y_t (top panel), jump probabilities (middle panel) and ex-post jump sizes (bottom panel), for the four exchange rates between January 3, 2005 and September 30, 2008. The algorithm identified numerous observations that have a high probability of being a jump. The average number of events with jump probability over 0.5 is roughly between 8 per year (for GBP) and 11 per year (for EUR), which is close to the values obtained from the EMM estimates over the full samples. The bimodal nature of the jump size distribution in the SVJD-B model guarantees that most of the identified jumps will be significant in size. This is an important feature of the model. For example, when the probability of a jump in the Euro exchange rate is greater than 0.5, the expected sizes fall within two bounds: the negative one, $[-1.45, -0.41]$ percent, and the positive one, $[0.53, 1.72]$ percent. In the usual SVJD specification with unimodal distribution of jump sizes most of the jumps are difficult to identify. This is because majority of them have a magnitude that is relatively close to the unconditional expectation, which is often very small.

Some jumps are isolated events, while others tend to cluster and lead to higher volatility and even more jumps. The highest concentration of jumps is in 2008, of which most coincide with the events related to the sub-prime mortgage crisis. Other jumps often coincide with the important news related to macroeconomy or asset markets. Consider, for example, the Euro exchange rate (Figure 2.3). Eight jumps happened on the dates when the U.S. Commerce Department issued reports about trade balance, unemployment levels, retail sales or GDP growth. Five jumps coincide with the announcements by the U.S. Federal Reserve or the European Central Bank regarding monetary policy, and two of them with important fiscal policy moves made by the U.S. Senate. Ten jumps coincide with unusually large stock market movements in the United

States or Europe, three with the unexpected earnings announcements by some of the major U.S. corporations, and one with the Société Générale $7 billion trading fraud. The strong co-movement of the currency market and the stock market is consistent with the findings of Cao (2001). These results, although far from being conclusive, reinforce the intuition based on Merton (1976, 1990) that jumps provide a mechanism through which unanticipated information about the most important determinants of the underlying process enter the market.

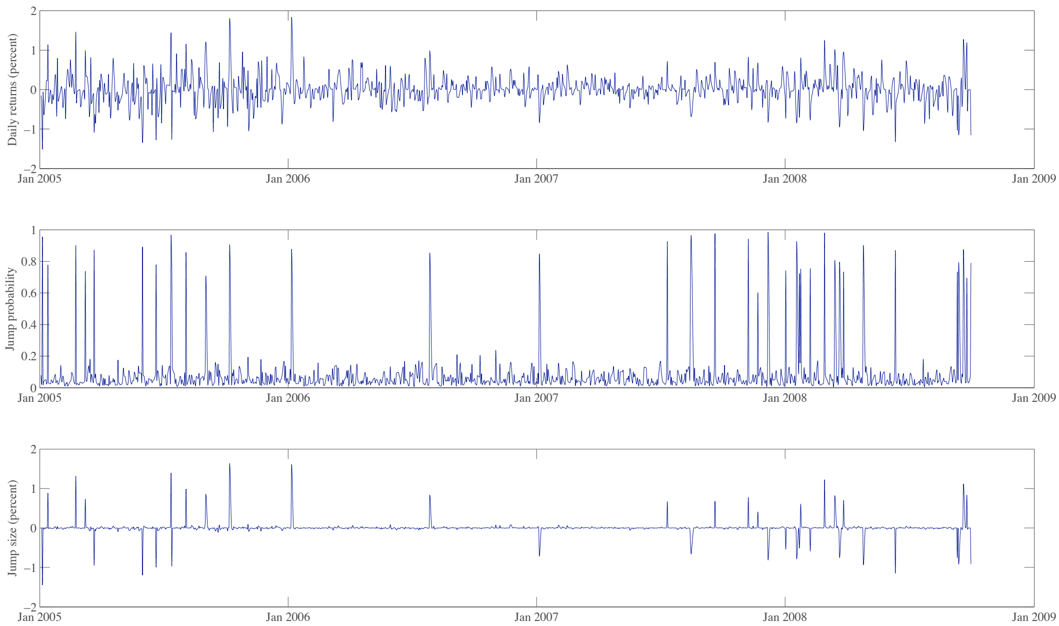

Figure 2.3: **EUR/USD exchange rate:** Returns, ex-post jump probabilities and expected jump sizes for January 2005–September 2008.

Figure 2.4: **GBP/USD exchange rate:** Returns, ex-post jump probabilities and expected jump sizes for January 2005–September 2008.

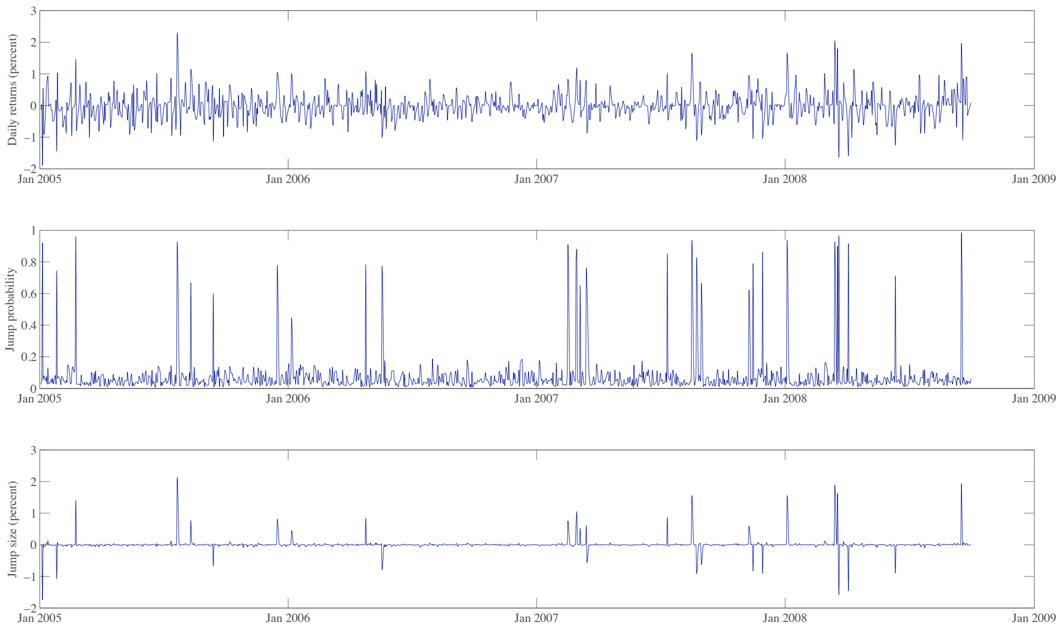

Figure 2.5: **JPY/USD exchange rate:** Returns, ex-post jump probabilities and expected jump sizes for January 2005–September 2008.

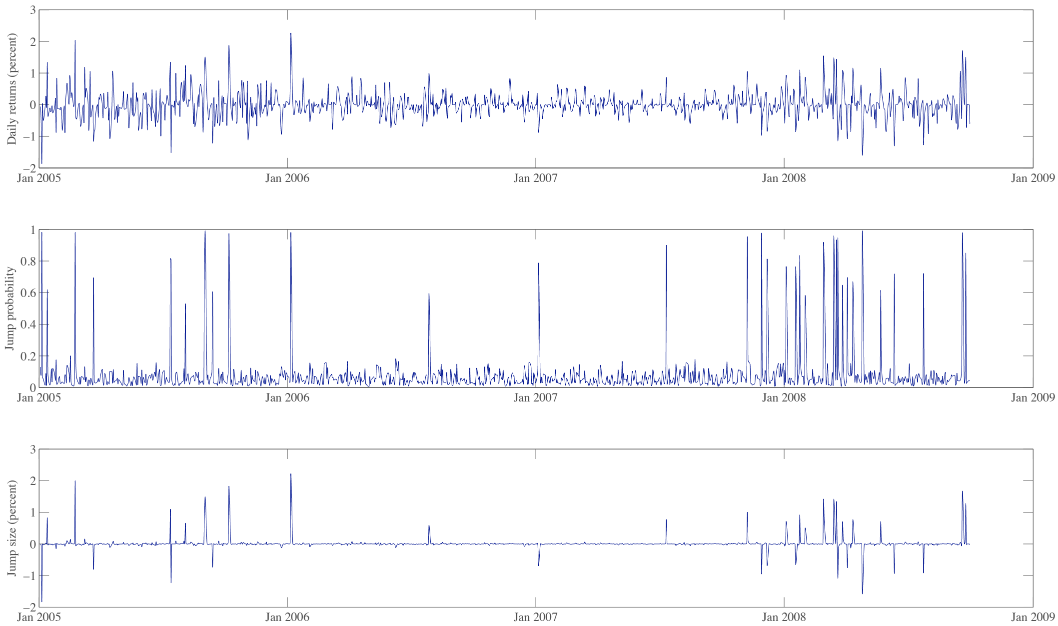

Figure 2.6: **CHF/USD exchange rate:** Returns, ex-post jump probabilities and expected jump sizes for January 2005–September 2008.

2.4 Option pricing implications

2.4.1 The impact of jumps on implied volatility patterns

The main empirical issue in option pricing is to find an appropriate model that will be consistent both with the observed dynamics of the underlying asset as well as with the observed option prices. The U-shaped patterns of implied volatilities, the so-called "volatility smiles", obtained from the actual data are difficult to reconcile with a great number of return models. This is also true for exchange rates, where any specification that does not allow for jumps fails to accommodate observed implied volatility patterns, even when stochastic nature of volatility is taken into account.

In this section I illustrate the effect of jumps on currency option prices. A suitable property of the SVJD-B model is that it can yield a closed-form solution for the price of European-style options. Options can be priced if the model specification is written in the risk-neutral form. Introducing the usual change of measure, the risk-neutral counterparts of the processes for the return and the instantaneous variance, equations (2.1) and (2.2), become

$$\begin{aligned} \frac{\mathrm{d}S_t}{S_t} &= \mu_t^* \mathrm{d}t + \sqrt{V_t}\,\mathrm{d}W_{1,t}^* + (e^{u_t} - 1)\,\mathrm{d}q_t - \lambda_t^* \bar{k}^* \mathrm{d}t, && (2.12) \\ \mathrm{d}V_t &= (\alpha - \beta_t^* V_t)\,\mathrm{d}t + \sigma\sqrt{V_t}\,\mathrm{d}W_{2,t}^*, && (2.13) \end{aligned}$$

where $\mu_t^* = r_t - r_t^f$ is the domestic-foreign interest rate differential. Stochastic processes $W_{1,t}^*$ and $W_{2,t}^*$ are now standard Brownian motions under the risk-neutral probability measure $\mathbb{P}^*$, having the same correlation coefficient as under the physical measure $\mathbb{P}$, that is $\mathrm{d}W_{1,t}^* \mathrm{d}W_{2,t}^* = \rho \mathrm{d}t$. The mean-reversion speed of the instantaneous variance β_t^* and the expected jump size $\lambda_t^* \bar{k}^* \mathrm{d}t$ depend on the market prices of volatility and jump risk, respectively. The explicit relationships are derived in Appendix A. The instantaneous risk premia are:

$$\begin{aligned} \text{premium for the return diffusion risk} &= \mu - \mu_t^*, \\ \text{premium for the volatility risk} &= (\beta - \beta_t^*)V_t, \\ \text{overall premium for the jump risk} &= \lambda_t \bar{k} - \lambda_t^* \bar{k}^*. \end{aligned}$$

The overall jump-risk premium consists of the combined premia for the uncertainty about the arrival of a jump and the uncertainty about the size of a jump.

At time t, the price of a European-style call option with the value of the underlying exchange rate equal to S_t, time to maturity τ and strike price X, is given by

$$C_t(S_t, V_t, \tau, X; \widehat{\psi}) = e^{-r_t^f \tau} S_t P_1(S_t, V_t, \tau, X; \widehat{\psi}) - e^{-r_t \tau} X P_2(S_t, V_t, \tau, X; \widehat{\psi}). \qquad (2.14)$$

Closed-form expressions for the functions P_1 and P_2 are given in Appendix B.

Various effects of stochastic volatility and jumps on option prices are illustrated in Figures 2.7–2.9. The graphs show generic examples, calculated for European-style call options on EUR/USD exchange rate. The curves represent the Black-Scholes implied volatilities

$$\sigma_{\text{imp}} = \text{BSImpVol}(S_t, C_t, r_t, r_t^f, \tau, X). \tag{2.15}$$

The implied volatilities σ_{imp} were obtained numerically, by substituting the values of C_t calculated with the formula (2.14) into equation (2.15). The set of parameters $\widehat{\psi}$ in (2.14) are the EMM estimates given in Table 2.6. The independent variable in Figures 2.7–2.9 is the relative moneyness, defined as the ratio of intrinsic value of option to the underlying exchange rate, i.e. $(S_t - X)/S_t$. All option prices C_t are computed for $S_t = 1.1512$, the sample average of the EUR/USD exchange rate. The U.S. and the Eurozone risk-free interest rates are set to $r_t = 0.02$ and $r_t^f = 0.05$, respectively. The instantaneous volatility $\sqrt{V_t}$ is fixed at the annualized long-run mean of 11.1433 percent.

Figure 2.7 displays the pricing effect of stochastic volatility and jumps, when there is no premium for volatility and jump risk ($\beta_t^* = \beta$, $\lambda_t^* = \lambda_t$ and $\bar{k}^* = \bar{k}$). The SV model produces a "smirk" pattern (dashed line), which is more pronounced for shorter maturities. This is indicative of a model in which the probability that the call option price will change significantly is low if the option is deep out of the money. The smirk effect wanes with maturity since the probability of moving towards higher prices increases with the remaining life of the option, while at the same time the probability of staying in the money decreases. In the SVJD-B model (full line), the jump component adds an upward tilt to the implied volatility, creating a familiar "smile" pattern. The smile virtually disappears at longer maturities. This effect has the following simple intuition. Jumps are not important for options with longer maturities, as they tend to be compensated in the long run. However, in the short run, the chance for a compensation is small. Therefore, jumps will make an impact on price as maturity date approaches: a deep-out-of-the-money option will have a non-negligible probability of ending up in the money only if the underlying exchange rate has a tendency to make sudden large jumps.

Figure 2.8 shows the effect of volatility risk premium implied by the SVJD-B model when jump risks premium is set to zero ($\lambda_t^* = \lambda_t$ and $\bar{k}^* = \bar{k}$). The instantaneous premium for volatility risk is measured by the difference between the speed of mean reversion β and its risk neutral counterpart β_t^*. I set the premium to 0 (full lines), 2 percent (dashed lines) and -2 percent (dotted lines). The graphs indicate that the volatility premium has little to no effect on short-maturity options. This is because unexpected changes of the underlying exchange rate over short time periods are mostly picked up by jumps, and if the jump risk premium is zero the exposure to the volatility risk alone has a negligible effect on option prices. At longer maturities, the exchange rate has more time to drift across the moneyness and hence the volatility risk becomes increasingly important. Positive premia decrease the long-run mean of the risk-neutral volatility, pushing the option prices down, and vice versa.

The impact of jump risk premium is shown in Figure 2.9. Now, the volatility premium implied by the SVJD-B model is set to zero ($\beta_t^* = \beta$), while the risk-neutral jump intensities take the values $\lambda_t^* = \lambda_t = 0.03$ (full line), $\lambda_t^* = 0.05$ (dashed line) and $\lambda_t^* = 0.07$ (dotted line). The risk-neutral expected jump size is set equal to its "physical" value, $\bar{k}^* = \bar{k} = 0.067$ percent. These values imply annual jump risk premia of 0, 0.5 and 1.0 percent, respectively. Even with relatively small premia, the effects are significant: a change in the risk-neutral jump intensity produces the twists in volatility smiles. The twists are more pronounced at short option maturities and show an asymmetric behavior. First, they are directed upward for out-of-the-money options and downward for in-the-money options. Second, the increase in implied volatility of out-of-the-money options is greater than the decrease of in-the-money options. A positive jump risk premium implies that the buyers require to be compensated for holding an option that is in the money to account for the risk of a negative jump. At the same time, they are willing to pay more for an out-of-the-money option, since higher jumps probabilities increase the chance to profit.

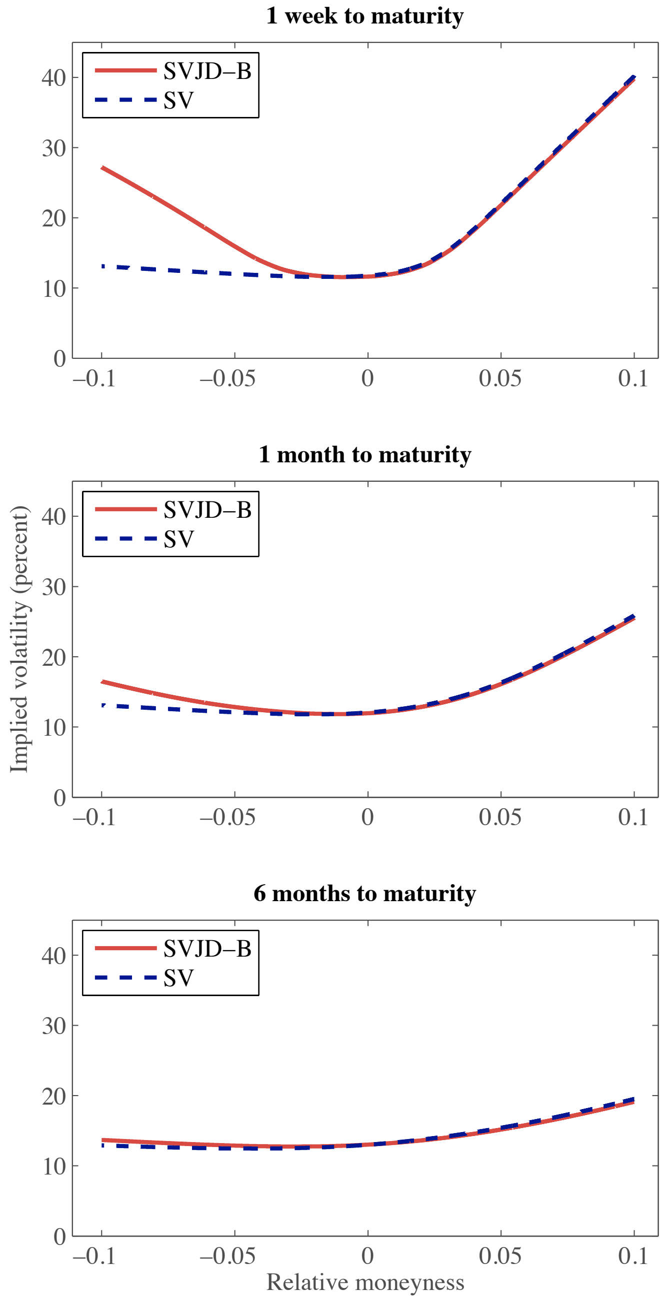

Figure 2.7: **The effect of stochastic volatility and jumps on option prices.** Black-Scholes implied volatilities are calculated from option prices generated by SVJD-B and SV models for the EUR/USD exchange rate. Model parameters are given in Table 2.6. The risk premia for the volatility and jump risks are set to zero. Panels display different times to maturity: 1 week, 1 month and 6 months.

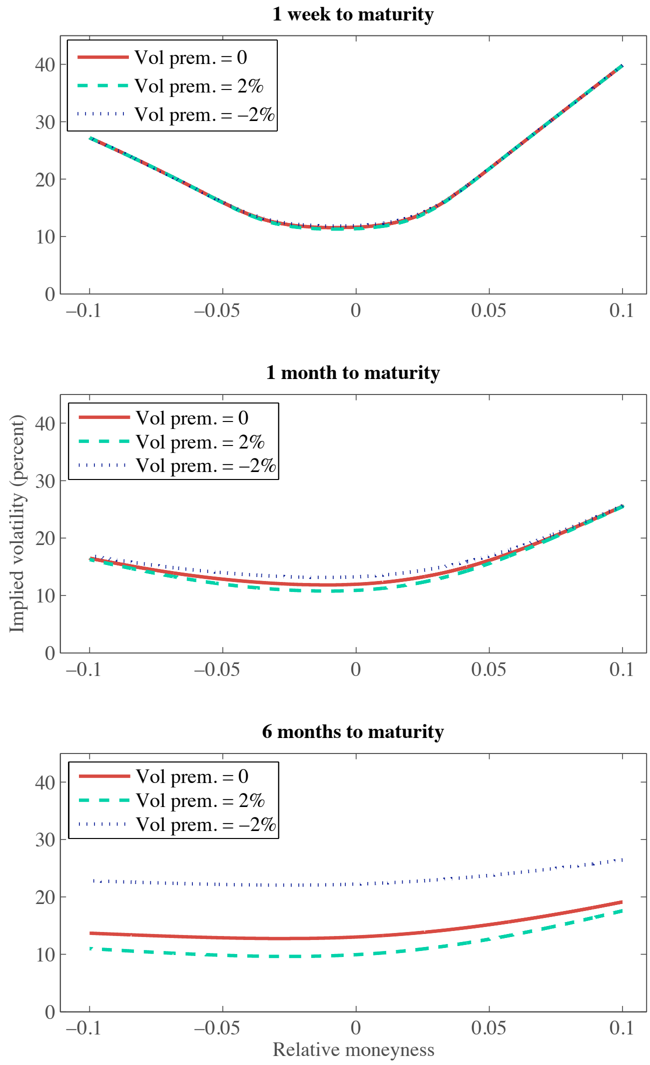

Figure 2.8: **The effect of volatility risk premium on option prices.** Black-Scholes implied volatilities are calculated from option prices generated by the SVJD-B model for the EUR/USD exchange rate. Model parameters are given in Table 2.6. Annual volatility risk premia are set to 0, 2 and −2 percent. Panels display different times to maturity: 1 week, 1 month and 6 months.

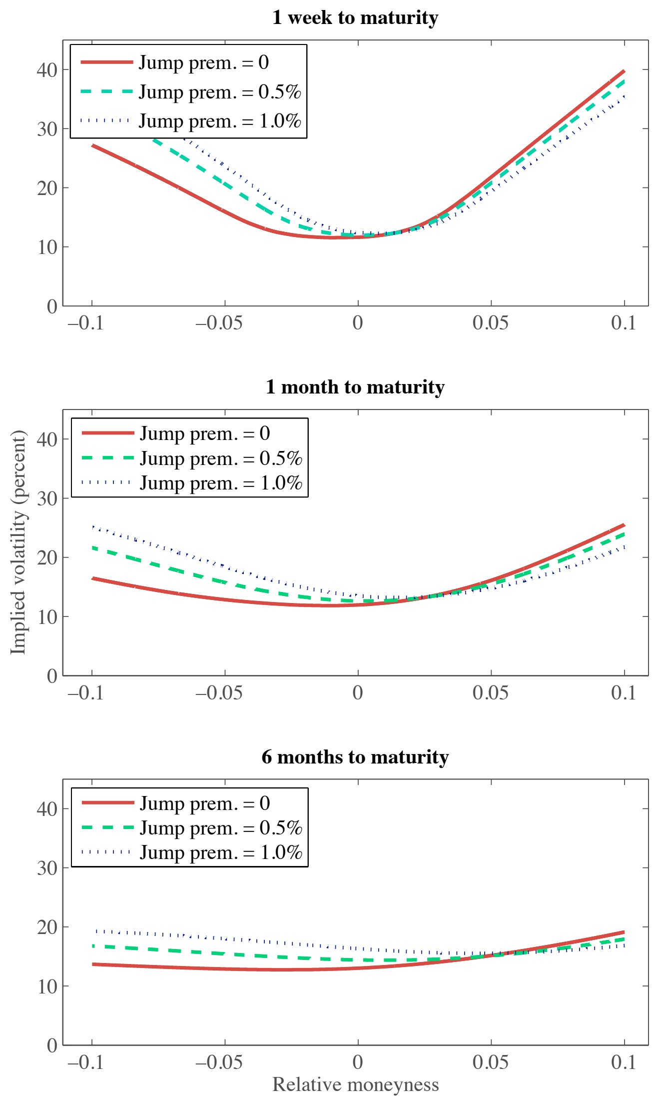

Figure 2.9: **The effect of jump risk premium on option prices.** Black-Scholes implied volatilities are calculated from option prices generated by the SVJD-B model for the EUR/USD exchange rate. Model parameters are given in Table 2.6. Annual jump risk premia are set to 0, 0.5 and 1.0 percent. Panels display different times to maturity: 1 week, 1 month and 6 months.

2.4.2 Risk premia and volatility smiles implicit in the cross-sectional currency options data

The SVJD-B model can fully accommodate the implied volatility patterns observed in the actual data. As an illustration, I use a cross section of European-style call options on Euro that were traded on the Philadelphia Stock Exchange (PHLX) on August 6, 2008. The PHLX currency options are settled in U.S. Dollars and expire on Saturday following the third Friday of the month. There were six available maturities: August 2008, September 2008, October 2008, December 2008, March 2009 and June 2009. The underlying exchange rate was $S_t = 1.5409$ and the available strikes went from 1.2700 to 1.6600, in steps of 0.0050, although some strike/maturity combinations had no open interest. There were 247 options in the cross section in total.

In order to match the model-implied options prices with the observed ones we need the risk-neutral parameter estimates. I use the yield on 3-month Treasury bill as a proxy for the U.S. risk-free rate and the 3-month Euribor as a proxy for the Eurozone risk-free rate. Their respective values on August 6, 2008 were $r_t = 1.4800$ percent and $r_t^f = 5.0289$ percent. Hence, the annualized risk-neutral drift rate was $\mu_t^* = -3.5489$ percent. This implies an annual premium for the return diffusion risk of 12.28 percent.

The remaining risk-neutral parameters, β_t^*, λ_0^*, λ_1^*, $\bar{k}^*$, as well as the instantaneous variance, V_t, can be obtained by solving

$$\begin{aligned} &\min_{\{V_t,\beta_t^*,\lambda_0^*,\lambda_1^*,\bar{k}^*\}} \sum_i w_i(\text{BSImpVol}_i^{\text{model}} - \text{BSImpVol}_i^{\text{data}})^2, \\ &w_i = \left(C_i^{\text{ask}} - C_i^{\text{bid}}\right)^{-1}. \end{aligned} \tag{2.16}$$

The estimator is designed to minimize the weighted squared difference between the Black-Scholes implied volatilities obtained from the data and the SVJD-B model. For every contract i, the point estimates of $\text{BSImpVol}_i^{\text{data}}$ are obtained from the average values of volatilities implied by the bid and the ask price.

To account for the differences in liquidity, the weights w_i are set equal to the reciprocal of the bid-ask spread of a given option contract. In this way, the contracts with higher liquidity will carry more weight in the estimation. The results of the optimization (2.16) are given in the left panel of Table 2.10. The Pearson's chi-square statistic indicates that the fit is highly significant. Figure 2.10 displays the market- and model-implied volatilities for four selected maturities. The error bars correspond to implied volatilities calculated from the bid and ask market prices, while the smooth lines are obtained from the SVJD-B model using the parameter estimates given in Tables 2.6 and 2.10. Parameter values imply annual risk premia of −2.30 and 0.16 percent for the volatility and jump risk, respectively (see the right panel of Table 2.10).

Table 2.10: **Option-implied parameters**
The left panel shows the instantaneous variance and risk-neutral parameters estimated from the cross section of currency option prices that were traded on PHLX on August 6, 2008. The right panel shows the corresponding risk premia.

Parameter	Value
V_t	0.0106
	(0.0015)
β_t^*	0.0248
	(0.0047)
λ_0^*	0.0332
	(0.0017)
λ_1^*	0.0027
	(0.0002)
$\bar{k}^*$	0.0007
	(0.0001)
$\chi^2[246]$	0.2973

(Standard errors in parentheses.)

Premium	Value (%)
Return diff. risk	12.28
	(4.98)
Volatility risk	−2.30
	(0.76)
Jump risk	0.16
	(0.03)

(Standard errors in parentheses.)

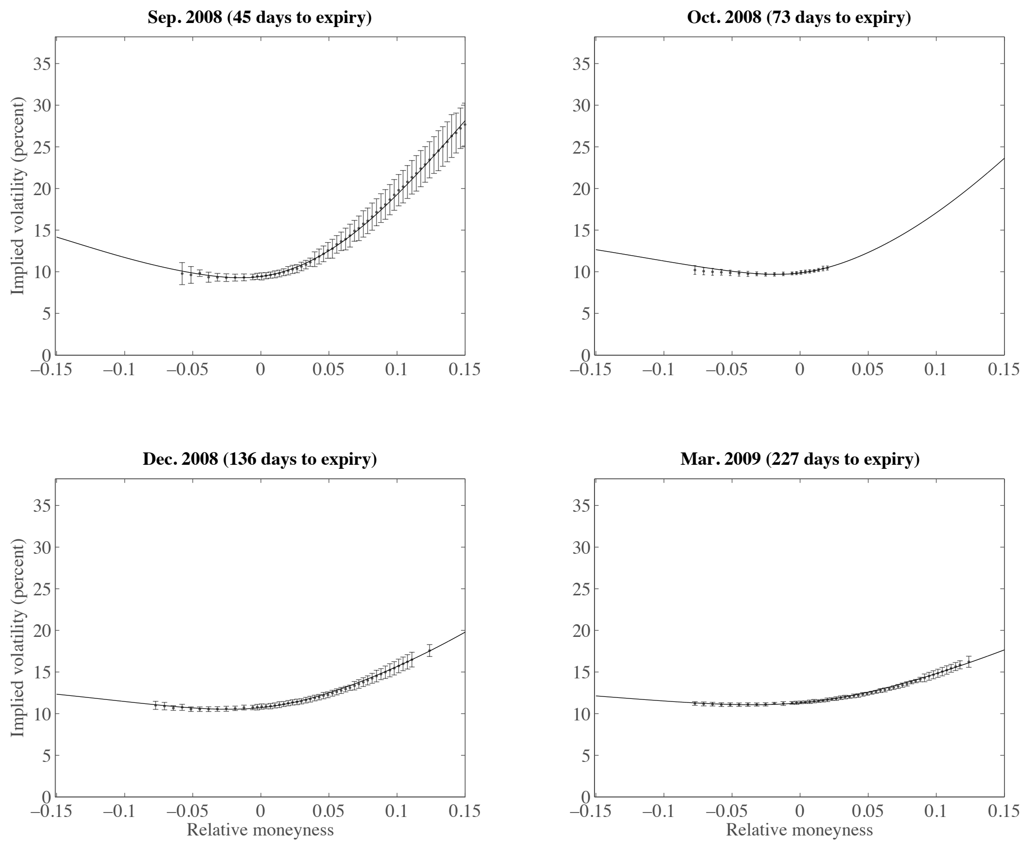

Figure 2.10: **Black-Scholes market- and model-implied volatilities.** Four selected maturities of European-style call option contracts on Euro. The error bars correspond to implied volatilities calculated from the bid and ask market prices quoted on PHLX on August 6, 2008. The smooth lines are obtained from the proposed SVJD-B model with instantaneous variance and risk-neutral parameters given in Table 2.10. Parameter values imply a volatility risk premium of -2.30 percent and a jump risk premium of 0.16 percent.

The premium for the return diffusion risk has the highest absolute value of the three, which is plausible given that the diffusion is responsible for most of the everyday changes. Volatility risk premium is negative and significant. The negative premium is a sign that investors are willing to pay more for exposure to the volatility uncertainty, which is reasonable given that higher volatility increases the option premium. The negative volatility risk premium is consistent with the findings of Bates (1996b). It is also implied in the prices of options on stock market indices (see, for example, Chernov and Ghysels (2000) or Pan (2002)). Finally, the jump risk premium is positive and significant, although an order of magnitude smaller than the volatility premium. Since jumps are very rare this is not surprising. However, the statistical significance of the jump risk premium indicates that the fear of jumps is important and seems to be priced by the market.

2.5 Conclusion

This chapter confirms the crucial role of stochastic volatility and jumps in exchange rate processes, at least in the four major U.S. Dollar-based spot exchange rates. The inference procedure based on the efficient method of moments shows that all pure-diffusion models are misspecified. These models are not able to capture the events in the tails of return distributions nor to accommodate the implied volatility patterns obtained from the actual options data. A stochastic volatility model with jump sizes from a bimodal distribution is able to fully remove the misspecification and yield an option pricing formula.

The filtering distributions of jump times inferred from the data indicate that jumps occur in irregular patterns, on average between eight and eleven times a year, depending on the exchange rate. In general, the jump probability weakly depends on volatility. On the other hand, jump events tend to coincide with the arrival of important news to the currency market. They also appear to be more frequent in the periods of turbulence in the stock market. This observation points to the importance of a deeper understanding of jumps in foreign exchange rates that goes beyond statistical significance.

Finally, jumps have a large impact on the prices of foreign currency options. They remove the distinct asymmetry of Black-Scholes implied volatility patterns characteristic for models without jumps. Moreover, the proposed general model is capable to accommodate the smile patterns observed in the actual data. Estimates of the risk-neutral model parameters obtained from the cross-sectional options data indicate that jump risk appears to be priced by the market.

Appendix A: The risk-neutral version of the model

Given that the diffusion and the jump process are independent of each other, we can split the return dynamics into the pure-diffusion part and the pure-jump part:

$$\frac{\mathrm{d}S_t}{S_t} = \left(\frac{\mathrm{d}S_t}{S_t}\right)_{\text{diff}} + \mathrm{d}J_t, \tag{2.17}$$

where

$$\left(\frac{\mathrm{d}S_t}{S_t}\right)_{\text{diff}} = \mu \mathrm{d}t + \sqrt{V_t}\ \mathrm{d}W_{1,t} \tag{2.18}$$

and

$$\mathrm{d}J_t = (e^{u_t} - 1)\,\mathrm{d}q_t - \lambda_t \bar{k} \mathrm{d}t. \tag{2.19}$$

Let us focus on the diffusion part first. Pure-diffusion return (2.18) and the instantaneous volatility V_t follow a joint Brownian diffusion, since W_1 and W_2 are correlated. Define

$$\mathrm{d}\mathbf{W}_t = \begin{bmatrix} \sqrt{V_t}\ \mathrm{d}W_{1,t} \\ \sigma\sqrt{V_t}\ \mathrm{d}W_{2,t} \end{bmatrix}, \tag{2.20}$$

for all t. To find the risk-neutral equivalent $\mathrm{d}\mathbf{W}^*$ of (2.20) that would be a martingale under an equivalent measure $\mathbb{P}^*$, we first write the Radon-Nikodým derivative of $\mathbb{P}^*$ with respect to the physical measure $\mathbb{P}$:

$$\frac{\mathrm{d}\mathbb{P}^*}{\mathrm{d}\mathbb{P}} = \exp\left[-\int_0^t \boldsymbol{\xi}_s \cdot \mathrm{d}\mathbf{W}_s - \frac{1}{2}\int_0^t \left(\boldsymbol{\xi}_s \cdot \mathrm{d}\mathbf{W}_s\right)\left(\mathrm{d}\mathbf{W}_s \cdot \boldsymbol{\xi}_s\right)\right],$$

where

$$\boldsymbol{\xi}_s = \begin{bmatrix} \xi_{1,s} \\ \xi_{2,s} \end{bmatrix}$$

is predictable at s (see Bingham and Kiesel (2004)). Then, by Girsanov's theorem, a $\mathbb{P}^*$-Brownian motion has the form

$$\mathrm{d}\mathbf{W}_t^* = \mathrm{d}\mathbf{W}_t \left(1 + \mathrm{d}\mathbf{W}_t \cdot \boldsymbol{\xi}_t\right).$$

Therefore,

$$\mathrm{d}\mathbf{W}_t \quad = \quad \mathrm{d}\mathbf{W}_t^* - \begin{bmatrix} 1 & \rho\sigma \\ \rho\sigma & \sigma^2 \end{bmatrix} \begin{bmatrix} \xi_{1,t} \\ \xi_{2,t} \end{bmatrix} V_t \mathrm{d}t,$$

which implies that we can substitute

$$\begin{bmatrix} \sqrt{V_t}\ \mathrm{d}W_{1,t} \\ \sigma\sqrt{V_t}\ \mathrm{d}W_{2,t} \end{bmatrix} = \begin{bmatrix} \sqrt{V_t}\ \mathrm{d}W_{1,t}^* - \left(\xi_{1,t} + \rho\sigma\xi_{2,t}\right) V_t \mathrm{d}t \\ \sigma\sqrt{V_t}\ \mathrm{d}W_{2,t}^* - \left(\rho\sigma\xi_{1,t} + \sigma^2\xi_{2,t}\right) V_t \mathrm{d}t \end{bmatrix}$$

into (2.1) and (2.2). Hence, the processes

$$\left(\frac{\mathrm{d}S_t}{S_t}\right)_{\text{diff}} = \mu_t^* \mathrm{d}t + \sqrt{V_t}\ \mathrm{d}W_{1,t}^*$$

and

$$\left(\alpha - \beta_t^* V_t\right) \mathrm{d}t + \sigma\sqrt{V_t}\ \mathrm{d}W_{2,t}^*$$

both contain diffusions that are (jointly) martingales under $\mathbb{P}^*$, as long as

$$\mu_t^* = \mu - \left(\xi_{1,t} + \rho\sigma\xi_{2,t}\right) V_t$$

and

$$\beta_t^* = \beta - \xi_t,$$

where $\xi_t \equiv \rho\sigma\xi_{1,t} + \sigma^2\xi_{2,t}$. The no-arbitrage argument in the form of covered interest parity requires that $\mu_t^* \mathrm{d}t = \mathbb{E}_t^*(\mathrm{d}S_t/S_t) = (r_t - r_t^f)\mathrm{d}t$. This constraint implies that at each t, $\xi_{1,t}$ and $\xi_{2,t}$ will not be independent given the values of the interest rates. A common assumption of constant elasticity of substitution in the utility function of the representative agent, as in Bates (1996b), will correspond to the case where ξ_t is constant in time.

The jump component in equation (2.19) is a $\mathbb{P}$-martingale by construction:

$$\begin{aligned}
\mathbb{E}_t(\mathrm{d}J_t) &= \mathbb{E}_t\left[(e^{u_t} - 1)\mathrm{d}q_t\right] - \lambda_t \bar{k}\mathrm{d}t \\
&= \mathbb{E}_t(e^{u_t} - 1)\lambda_t\mathrm{d}t - \lambda_t\bar{k}\mathrm{d}t \\
&= 0.
\end{aligned}$$

The second equality follows from measurability of V_t with respect to $\mathcal{F}_t$. Define

$$\mathrm{d}N_t = \mathrm{d}q_t - \lambda_t\mathrm{d}t.$$

By applying Girsanov's theorem for point processes (Elliot and Kopp (2005)), the risk-neutral version of $\mathrm{d}N$ will be

$$\begin{aligned}
\mathrm{d}N_t^* &= \mathrm{d}N_t - \mathbb{E}_t\left[\frac{e^{a+bu_t}}{\mathbb{E}_t(e^{bu_t})} - 1\right]\lambda_t\mathrm{d}t \\
&= \mathrm{d}N_t - (e^a - 1)\lambda_t\mathrm{d}t \\
&= \mathrm{d}q_t - \lambda_t^*\mathrm{d}t,
\end{aligned}$$

where the market prices of jump risk a and b are measurable with respect to $\mathcal{F}_t$, and $\lambda_t^* \equiv e^a\lambda_t$. Girsanov's theorem applied to $\mathrm{d}J$ then yields

$$\begin{aligned}
\mathrm{d}J_t^* &= \mathrm{d}J_t - \mathbb{E}_t\left[\left(e^a\frac{e^{bu_t}}{\mathbb{E}_t(e^{bu_t})} - 1\right)(e^{u_t} - 1)\right]\lambda_t\mathrm{d}t \\
&= (e^{u_t} - 1)\mathrm{d}q_t - \lambda_t^*\left[e^{b\omega^2}\frac{Q(b+1)}{Q(b)} - 1\right]\mathrm{d}t,
\end{aligned}$$

where

$$Q(\Phi) = p(1+k)^{\Phi} + (1-p)(1-k)^{\Phi}.$$

Therefore, the process

$$\mathrm{d}J^* = (e^{u_t} - 1)\mathrm{d}q_t - \lambda_t^* \bar{k}^* \mathrm{d}t$$

will be a martingale under $\mathbb{P}^*$ as long as

$$\bar{k}^* = e^{b\omega^2} \frac{Q(b+1)}{Q(b)} - 1.$$

Parameter a captures the inability of the market to time the arrival of jumps, while b measures the uncertainty related to the jump size and, possibly, the model uncertainty. Liu, Pan, and Wang (2005) also argue that a significant part of the jump risk premium should come from the uncertainty aversion in the sense of Knight (1921) and Ellsberg (1961).

Putting everything together, the processes

$$\begin{aligned}
\frac{\mathrm{d}S_t}{S_t} &= \mu_t^* \mathrm{d}t + \sqrt{V_t}\mathrm{d}W_{1,t}^* + (e^{u_t} - 1)\,\mathrm{d}q_t - \lambda_t^* \bar{k}^* \mathrm{d}t, \\
\mathrm{d}V_t &= (\alpha - \beta_t^* V_t)\,\mathrm{d}t + \sigma\sqrt{V_t}\mathrm{d}W_{2,t}^*,
\end{aligned}$$

with $\mathrm{d}W_{1,t}^* \mathrm{d}W_{2,t}^* = \rho \mathrm{d}t$, represent the risk-neutral equivalents of (2.1) and (2.2). The market risk premia are the following:

$$\begin{aligned}
\text{premium for the return diffusion risk} &= \mu - \mu_t^* \\
\text{premium for the volatility risk} &= (\beta - \beta_t^*)V_t \;=\; -\xi_t V_t \\
\text{overall premium for the jump risk} &= \lambda_t \bar{k} - \lambda_t^* \bar{k}^*.
\end{aligned}$$

Appendix B: Closed-form solution for the price of a European currency option

Given the risk-adjusted model (2.12)–(2.13), the price at t of a European call option with residual maturity $\tau = T - t$ and strike price X is given by

$$
\begin{aligned}
C_t(S_t, V_t, \tau, X; \widehat{\psi}) &= e^{-r_t \tau} \mathbb{E}_t^* \left[\max \left(S_T - X, 0 \right) \right] \\
&= e^{-r_t^f \tau} S_t P_1 - e^{-r_t \tau} X P_2,
\end{aligned}
$$

where $\mathbb{E}_t^*(\cdot)$ denotes the expectation with respect to the risk-neutral probability measure $\mathbb{P}^*$ and conditional on the sigma-algebra $\mathcal{F}_t$. P_1 and P_2 have the usual Black-Scholes interpretation of the expected value of the underlying asset conditionally on the option being in the money, and probability of being in the money, respectively. The closed-form expressions for P_1 and P_2 can be obtained by following the calculation steps similar to those in Bates (1996b). The results are

$$
\begin{aligned}
P_j &= \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \frac{\operatorname{imag} \left(F_j(i\Phi) e^{-i\Phi x} \right)}{\Phi} \mathrm{d}\Phi, \\
F_j(\Phi; V, \tau) &= \exp \left\{ A_j(\tau; \Phi) + B_j(\tau; \Phi) V + \lambda_0^* \tau C_j(\Phi) \right\}, \\
A_j(\tau; \Phi) &= \mu_t^* \tau \Phi - \frac{\alpha \tau}{\sigma^2} (\rho \sigma \Phi - \beta_j - \gamma_j) \\
&\quad - \frac{2\alpha}{\sigma^2} \ln \left[1 + \frac{1}{2} (\rho \sigma \Phi - \beta_j - \gamma_j) \frac{1 - e^{\gamma_j \tau}}{\gamma_j} \right], \\
B_j(\tau; \Phi) &= - \frac{\Phi^2 + (3 - 2j)\Phi + 2\lambda_1^* C_j(\Phi)}{\rho \sigma \Phi - \beta_j + \gamma_j (1 + e^{\gamma_j \tau}) / (1 - e^{\gamma_j \tau})}, \\
C_j(\Phi) &= (1 + \bar{k}^*)^{2-j} \left[Q(\Phi; p, k^*) e^{(1/2)\omega^2 (\Phi^2 + (3-2j)\Phi)} - 1 \right] - \bar{k}^* \Phi, \\
\gamma_j &= \sqrt{(\rho \sigma \Phi - \beta_j)^2 - \sigma^2 \left[\Phi^2 + (3 - 2j)\Phi + 2\lambda_1^* C_j(\Phi) \right]}, \\
\beta_j &= \beta_t^* + \rho \sigma (j - 2), \\
Q(\Phi; p, k^*) &= p(1 + k)^\Phi + (1 - p)(1 - k)^\Phi, \\
x &= \ln(X / S_t),
\end{aligned}
$$

$$\mu_t^* = r_t - r_t^f,$$

for $j = 1, 2$. By setting $p = 1$ and $\lambda_1^* = 0$ we obtain the option pricing formula given in Bates (1996b) for currency options, or in Bates (2000) for options on a stock market index.

Chapter 3

An Efficient Method for Market Risk Management under Multivariate Extreme Value Theory Approach

3.1 Introduction

Sometimes extreme times indeed call for extreme measures. Events like financial crises and market crashes have increased awareness of the need to quantify risk and assess the probability and extent of extremely large losses. Currently, the most popular tool used by financial institutions to measure and manage market risk is Value at Risk (VaR). VaR refers to the maximum potential loss over a given period at a certain confidence level. Originally intended as a reporting tool for senior management, it started becoming prevalent in the risk management world in 1994, when JPMorgan published the methodology behind its *RiskMetrics* system. Soon after, books by Jorion (1996) and Dowd (1998) introduced VaR to academic parlance and gave it more formal theoretical ground. VaR quickly entered other core areas of banking such as capital allocation, portfolio optimization or risk limitation. With its increasing importance, VaR was easily adopted by the regulators as well. In particular, the Basel II capital requirements for market risk are based on VaR.

In spite of being established an industry and regulatory standard, VaR is often criticized for not being a coherent risk measure.[1] Namely, VaR is not strictly sub-aditive, since there might be situations in which $\mathrm{VaR}(X+Y) > \mathrm{VaR}(X) + \mathrm{VaR}(Y)$, as shown for example in Artzner, Delbaen, Eber, and Heath (1999), Acerbi and Tasche (2002) or Breuer, Jandacka, Rheinberger, and Summer (2008). Furthermore, VaR completely ignores statistical properties of losses beyond the specified quantile of the profit-loss distribution, i.e. the tail risk. In order to overcome these drawbacks, Artzner, Delbaen, Eber, and Heath (1997) proposed the Expected Shortfall (ES) as an alternative risk measure. It is defined as the conditional expectation of loss beyond a fixed level of VaR. As such, ES takes into account tail risk and satisfies the sub-aditivity property, which assures its coherence as a risk measure.

VaR and ES are usually estimated in analytical, simulation or historical framework. Analytical approach relies upon the assumption that returns or return innovations follow a known distribution, such as normal. Since financial time series commonly exhibit significant autocorrelation and heteroskedasticity, one typically models the conditional rather than the unconditional distribution of returns. However, many empirical results, such as McNeil (1997), da Silva and de Melo Mendez (2003) and Jondeau and Rockinger (2003), show that the normality assumption fails in explaining extreme events, even when autocorrelation and heteroskedasticity are taken into account. This follows from the fact that the high-frequency empirical returns are characterized by heavier tails than those implied by the normal distribution, as well as by a substantial skewness. In order to overcome these problems, a leptokurtic and/or skewed distribution, such as (standard or skewed) Student's t, may be used instead. However, empirical results based on the t-distribution have shown

[1] A coherent risk measure satisfies properties of monotonicity, sub-additivity, homogeneity and translational invariance.

only a limited success. Alternatively, we can estimate VaR and ES via a simulation. The simulation method is quite useful, if not the only one available, when the underlying risk factors have non-linear payoffs, which is the case with options, for example. However, any simulation has to be based on a prespecified model of dynamics of the underlying factors, thus the VaR and ES estimates will critically rely on a correct model specification with properly and precisely calibrated parameters.

To avoid ad-hoc assumptions of (un)conditional return distribution or dynamics of the underlying risk factors, the historical simulation (HS) is often used as an alternative. The HS employs historical data from recent past, thereby allowing for the presence of heavy tails without making assumptions about the probability distribution or dynamics of returns. This non-parametric approach is conceptually simple as well as easy to implement. Moreover, it entirely overcomes the problem of model risk. Unfortunately, it suffers from some serious drawbacks. First, any extrapolation beyond past observations will be inaccurate, especially if the historical series is relatively short. If we try to mitigate this problem by considering longer samples, we will practically neglect the time-varying nature of volatility, as well as volatility clustering. In that case, the HS approach would not properly capture the risk in a sudden period of extremely high volatility – the VaR and ES estimates would change only marginally.

Beyond these traditional approaches, there is an alternative which uses the Extreme Value Theory (EVT) to characterize the tail behavior of the distribution of returns. By focusing on extreme losses, the EVT successfully avoids tying the analysis down to a single parametric family fitted to the whole distribution. Although there is a history of use of EVT in the insurance industry, its application to market risk calculations began about a decade ago. McNeil (1999), Bensalah (2000), Smith (2000), Nyström and Skoglund (2002b) and Embrechts, Klüppelberg, and Mikosch (2008) survey the mathematical foundations of EVT and discuss its applications to both financial and insurance risk management. The empirical results show that EVT-based models provide more accurate VaR estimates, especially in higher quantiles. For example, McNeil (1997), Nyström and Skoglund (2002b), Harmantzis, Chien, and Miao (2005) and Marinelli, d'Addona, and Rachev (2007) show that EVT outperforms the estimates of VaR and ES based on analytical and historical methods.

EVT approach thus seems like a natural choice for risk measurement: its implementation is relatively easy and is based on a few assumptions required for the asymptotics to work. Regrettably, this elegance comes with a price, as the straightforwardness is limited to the univariate EVT. In practice, the number of assets in a typical portfolio is large. We usually deal with a multitude of risk factors and hence our measurement method requires a multivariate approach. However, defining a multivariate model for the evolution of risk factors under extreme market conditions has so far been a daunting task. A seemingly obvious technique involves a multivariate version of the EVT, based on the multidimensional limiting relations (see Smith (2000)), but model complexity increases greatly with the number of risk factors. Alternatively, the joint distribution of returns can be seen as a product of marginal distributions and a copula. McNeil and Frey (2000) and Nyström and Skoglund (2002a), for example, describe the copula approach to assessment of the extreme co-dependence structure of risk factors. Not only that this technique introduces an additional model risk, inherent in the assumption of a specific analytical form of the co-dependence function, but it also becomes quite intractable with increase in dimensionality. Moreover, a typical copula method for multivariate EVT, such as the one described in Nyström and Skoglund (2002a), requires an additional simulation step in order to retrieve the innovations from the joint distribution, given the fitted marginals and parameters of the copula.

This chapter introduces a multivariate EVT method for risk measurement that is based on separate estimations of the univariate model. A key assumption of the univariate EVT is that extreme returns are independent and identically distributed. Instead of estimating the joint n-dimensional distributions (using copulas or otherwise), the proposed method works with n orthogonal series of conditional residuals that are approximately independent and identically distributed. These residuals are obtained from the principal components of the joint return series that are free of any autocorrelation, heteroskedasticity and asymmetry. The latter is achieved by assuming that the joint return process follows a stationary n-dimensional model from the ARMA-GARCH family. To render the method free of any unnecessary distributional assumption, the ARMA-GARCH parameters are estimated by a generalized method of moments.

As an illustration, the technique is applied to a sequence of daily interbank spot exchange rates of Euro, British Pound, Japanese Yen and Swiss Franc with respect to the U.S. Dollar. The VaR and ES estimates are compared to the actual losses. The results indicate that the method performs well in jointly capturing extreme events in all four series. It also yields more precise VaR and ES estimates and forecasts than the usual methods based on conditional normality, conditional t-distribution or historical simulation.

The remainder of the chapter is organized as follows: Section 3.2 presents the theoretical background behind the EVT approach and the estimation methodology used in this chapter. Section 3.3 describes the data and provides an example of estimation. Section 3.4 shows the back-tests of the model and its forecasting ability, and compares these results to the ones corresponding to the usual methods applied in risk modeling. Concluding remarks are given in Section 3.5.

3.2 Theoretical Framework and Estimation Methodology

3.2.1 Theoretical Framework

This subsection outlines some basic results of the univariate extreme value theory. First, I formally define the two risk measures used throughout the chapter, VaR and ES. Next, I present two most important results of EVT that concern the asymptotic distributions of the order statistics and of the exceedances over a given treshold.

Definition 1 *Let* $\{X_i\}_{i=1}^n$ *be a set of independent and identically distributed random variables with distribution function*

$$F(x) := \mathbb{P}\{X_i \leq x\}$$

for any i. Value at Risk *is the* q*-th quantile of the distribution* F*:*

$$\mathrm{VaR}_q := F^{-1}(q),$$

where $q \in (0,1)$ *and* F^{-1} *is the inverse of* F*. Similarly, the* Expected Shortfall *is the expected value of* X*, given that* VaR *is exceeded:*

$$\mathrm{ES}_q := \mathbb{E}[X|X > \mathrm{VaR}_q].$$

In order to compute VaR and ES we have to be able to assess the upper and lower tails of the distribution function F. Hence, it is natural to consider the order statistics

$$\begin{aligned} M_n &= \max\{X_1, X_2, \ldots, X_n\}, \\ m_n &= \min\{X_1, X_2, \ldots, X_n\}. \end{aligned}$$

Both M_n and m_n are random variables that depend on the length n of the sample. In analogy with the Central Limit Theorem, we will be interested in the asymptotic behavior of these random variables as $n \to \infty$. Since $m_n = -\max\{-X_1, -X_2, \ldots, -X_n\}$ it is sufficient to state all the results for M_n, that is, focus on observations in the upper tail of the underlying distribution. The results for the lower tail will be straightforward to generalize.

The following theorem is a limit law first derived heuristically by Fisher and Tippett (1928) and later from a rigorous standpoint by Gnedenko (1943).

Theorem 1 *Let $\{X_i\}_{i=1}^n$ be a set of n independent and identically distributed random variables with distribution function F and suppose that there are sequences of normalization constants, $\{a_n\}$ and $\{b_n\}$, such that, for some non-degenerated limit distribution F^*, we have*

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{M_n - b_n}{a_n} \le x\right) = \lim_{n\to\infty} [F(a_n x + b_n)]^n = F^*(x), \quad x \in \mathbb{R}.$$

Then, there exist $\xi \in \mathbb{R}$, $\mu \in \mathbb{R}$ and $\sigma \in \mathbb{R}_+$ such that $F^*(x) = \Gamma_{\xi,\mu,\sigma}(x)$ for any $x \in \mathbb{R}$, where

$$\Gamma_{\xi,\mu,\sigma}(x) := \exp\left[-\left(1 + \xi\frac{x-\mu}{\sigma}\right)_+^{-1/\xi}\right]$$

is the so-called generalized extreme value (GEV) *distribution.*

The GEV was first proposed by von Mises (1936) in this form. The $1/\xi$ is referred to as the tail index, as it indicates how heavy the upper tail of the underlying distribution F is. When $\xi \to 0$, the tail index tends to infinity and $\Gamma_{\xi,\mu,\sigma}(x) \to \exp\left[-\exp\left(-(x-\mu)/\sigma\right)\right]$.

The sign of ξ defines the three fundamental types of extreme value distributions:

• If $\xi = 0$, the distribution is called the Gumbel distribution. In this case, the distribution spreads out along the entire real axis.

• If $\xi > 0$, the distribution is called the Fréchet distribution. In this case, the distribution has a lower bound.
• If $\xi < 0$, the distribution is called the Weibull distribution. In this case, the distribution has an upper bound.
Many of the well known distributions may be divided between these three classes of GEV distribution according to their behavior in the tail. For example, normal, gamma and log-normal distributions converge to Gumbell distribution ($\xi = 0$); Student's t, Pareto, log-gamma and Cauchy converge to Fréchet distribution ($\xi > 0$); uniform and beta converge to Weibull distribution ($\xi < 0$). The subset of all distributions F that converge to a given type of extreme value distribution is called the *domain of attraction* for that type. Some characterizations of a domain of attraction are given in Nyström and Skoglund (2002b). More details on GEV distribution and domains of attraction can be found, for example, in Embrechts, Klüppelberg, and Mikosch (2008).

EVT is sometimes applied directly – for example, by fitting GEV to the maxima of the series, see Smith (2000). An alternative approach is based on *exceedances over treshold.* The following theorem, first stated by Picklands (1975), gives the asymptotic form of conditional distribution beyond a very high treshold.

Theorem 2 *Let $\{X_i\}_{i=1}^n$ be a set of n independent and identically distributed random variables with distribution function F. Define*

$$F_u(y) := \mathbb{P}\left(X \leq u + y \mid X > u\right) = \frac{F(u+y) - F(u)}{1 - F(u)}, \quad y > 0$$

to be the distribution of excesses of X over the treshold u. Let x_F be the end of the upper tail of F, possibly a positive infinity. Then, if F is such that the limit given by Theorem 1 exists, there are constants $\xi \in \mathbb{R}$ and $\beta \in \mathbb{R}_+$ such that

$$\lim_{u \to x_F} \sup_{u < x < x_F} |F_u(x) - G_{\xi,\beta}(x - u)| = 0,$$

where

$$G_{\xi,\beta}(y) := 1 - \left(1 + \xi\frac{y}{\beta}\right)_+^{-1/\xi} \tag{3.1}$$

is known as the generalized Pareto (GP) *distribution.*

There is a close analogy between Theorems 1 and 2 because ξ is the same in both, and there is a one-to-one correspondence between GEV and GP distributions, given by

$$1 - G_{\xi,\beta}(x) = -\ln \Gamma_{\xi,0,\sigma}(x),$$

see Balkema and de Haan (1974), Davison and Smith (1990) and Nyström and Skoglund (2002b).

The application of EVT involves a number of challenges. First, the parameter estimates of the GEV and GP limit distributions will depend on the number of extreme observations used. Second, the choice of a treshold should be large enough to satisfy the conditions that permit the application of Theorem 2, i.e. $u \to x_F$, while at the same time leaving a sufficient number of observations to render the estimation feasible. There are different methods of making this choice, and some of them are examined in Bensalah (2000). Finally, Theorems 1 and 2 hold only if the extreme observations X are independent and identically distributed. Therefore, we cannot apply the results of EVT to returns on financial assets directly, since a typical financial time series exhibits autocorrelation and heteroskedasticity. Moreover, the EVT approach described in this subsection applies only to a single time series, whereas in practice we often deal with multidimensional series. The following subsection describes how to overcome these issues.

3.2.2 Estimation Methodology

Estimating Independent Univariate Excess Distributions

Theorem 2 states that for a large class of underlying excess distributions (namely, those satisfying Theorem 1), the distribution of exceedances over treshold converges to a generalized Pareto as the treshold is raised. Thus, the GP distribution is the natural model for the unknown excess distribution. The excess distribution above the threshold u may be therefore taken to be *exactly* GP for some ξ and β:

$$F_u(y) = G_{\xi,\beta}(y), \tag{3.2}$$

for any y satisfying $0 \leq y < x_F - u$.

Assuming that we have a set of realizations $\{z_{t,i}\}_{t=1}^{T}$, we can choose a sensible treshold u and estimate parameters ξ and β. If there are N_u out of a total of T data points that exceed the threshold, the GP will be fitted to the N_u exceedances. In the literature, several estimators have been used to fit the parameters of the GP distribution. Two most popular ones are the maximum likelihood (ML) and the Hill estimator. The ML estimator is based on the assumption that if the tail under consideration exactly follows a GP distribution, then the likelihood function can be written in a closed form. The estimators of the parameters ξ and β are then obtained using the standard ML approach. Provided that $\xi > -1/2$ the ML estimator of the parameters is consistent and asymptotically normal as the number of data points tends to infinity. The alternative is based on a combination of the ML method and the following semi-parametric result.

Theorem 3 *Suppose $\{X_t\}_{t=1}^{T}$ are independent and identically distributed random variables with distribution function F, and*

$$\lim_{k\to\infty} \frac{1-F(kx)}{1-F(k)} = x^{-1/\xi}, \quad x \in \mathbb{R}_+, \quad \xi > 0.$$

Then, for $x > 0$,

$$\lim_{T\to\infty} \mathbb{P}\left(\frac{M_T - b_T}{a_T} \leq x\right) = \Gamma_{\xi,0,1}(x),$$

where $b_T = 0$ and $a_T = F(1 - 1/T)$.

When estimating ξ one may, assuming a priori that $\xi > 0$, conjecture that the tail behaves as in Theorem 3 and obtain an ML estimator of the parameter ξ. This estimator is referred to as the Hill estimator, see Danielsson and de Vries (1997).

Nyström and Skoglund (2002b) have shown that ML typically performs better than the Hill estimator in terms of relative bias and relative standard deviation. In addition, ML has a useful property of being almost invariant to the choice of threshold. This is in sharp contrast to the Hill estimator which is very sensitive to this choice. Also, the Hill estimator is designed specifically for the heavy-tailed case whereas the ML method is applicable to light-tailed data as well.

Estimating Tails of Univariate Distributions

By setting $x = u + y$ and combining Theorem 2 and expression (3.2) we can write

$$F(x) = (1 - F(u))\, G_{\xi,\beta}(x - u) + F(u),$$

for $x > u$. This formula shows that we may easily interpret the model in terms of the tail of the underlying distribution $F(x)$ for $x > u$. Thus, the only additional element we require to construct a tail estimator is an estimate of $F(u)$. For this purpose, I use the method of historical simulation (HS) and take the obvious empirical estimator, $\widehat{F}(u) = 1 - N_u/T$. By setting a threshold at u, we are assuming that we have sufficient observations exceeding u for a reasonable HS estimate of $F(u)$, but for observations beyond u the historical method would be unreliable. Alternatively, we can find N_u that is closest to a predetermined $F(u)$. Thus, for example, in a sample of $T = 1000$ observations, $\widehat{F}(u) = 0.90$ will correspond to $N_u = 100$. The treshold is then set to $u = X_{900}$, if $\{X_t\}_{t=1}^T$ are ordered from the lowest to the highest.

Combining the HS estimate $\widehat{F}(u)$ with the ML estimates of the GP parameters, we obtain the tail estimator:

$$\widehat{F}(x) = 1 - \frac{N_u}{T}\left(1 + \widehat{\xi}\frac{x-u}{\widehat{\beta}}\right)^{-1/\widehat{\xi}}, \quad x > u. \tag{3.3}$$

Note that when the scale parameter β tends to infinity, $G_{\xi,\beta}(\cdot)$ vanishes and the tail estimator converges to the empirical one for any x. Thus, the tail estimator in (3.3) can be viewed as the HS estimator augmented by the tail behavior, which is captured by the GP distribution.

Estimating Univariate VaR and ES

For a given upper-tail probability $q > F(u)$ the VaR estimate is calculated by inverting the tail estimation formula (3.3) to get

$$\widehat{\mathrm{VaR}}_q = u + \frac{\widehat{\beta}}{\widehat{\xi}}\left[\left(\frac{T}{N_u}(1-q)\right)^{-\widehat{\xi}} - 1\right]. \tag{3.4}$$

This is a quantile estimate, where the quantile is an unknown parameter of an unknown underlying distribution. The confidence interval for $\widehat{\mathrm{VaR}}_q$ can be obtained using a method known as the *profile likelihood.*

Once we have $\widehat{\mathrm{VaR}}_q$, the point estimator of ES can be obtained from

$$\widehat{\mathrm{ES}}_q = \frac{1}{1-\widehat{\xi}}\left(\widehat{\mathrm{VaR}}_q + \widehat{\beta} - \widehat{\xi}u\right), \tag{3.5}$$

(see, for example, McNeil (1999)). As the tail index increases (equivalently, as $\widehat{\xi} \to 0$), the ES becomes progressively greater than VaR.

It is now easy to generalize the results for the VaR and ES such that they hold in the lower tail as well. Let $u_+ \equiv u$ be the upper-tail treshold, and let the lower-tail treshold u_- be defined symmetrically, that is by $F(u_-) = 1 - F(u_+)$. Then, for a given upper-tail probability $q_+ > F(u_+)$ or a given lower-tail probability $q_- < F(u_-)$ the general form of the VaR estimate is

$$\widehat{\mathrm{VaR}}_{q_\pm} = u_\pm \pm \frac{\widehat{\beta}_\pm}{\widehat{\xi}_\pm} \left[\left(\frac{T}{N_{u_\pm}} (1 - q_\pm) \right)^{-\widehat{\xi}_\pm} - 1 \right], \tag{3.6}$$

where the subscript $+$ $(-)$ refers to parameters in the upper (lower) tail. Similarly, the general form of the ES estimate is

$$\widehat{\mathrm{ES}}_{q_\pm} = \frac{1}{1 - \widehat{\xi}_\pm} \left(\widehat{\mathrm{VaR}}_{q_\pm} \pm \widehat{\beta}_\pm - \widehat{\xi}_\pm u_\pm \right). \tag{3.7}$$

It is important to stress that the interpretation of VaR and ES may vary, depending on the meaning of the set of variables $\{X_t\}_{t=1}^T$. Usually, in the risk modeling context these variables represent profits and hence are expressed in monetary units.

Orthogonalization

The ultimate goal is to apply the EVT approach to a portfolio consisting of n assets. Before we can use any of the results of EVT outlined in Subsection 3.2.1, we have to construct a set of cross-sectionally uncorrelated random variables. A natural choice is to work with the principal components of the unconditional covariance matrix of the log returns.

Definition 2 *Define $\boldsymbol{\varepsilon}_t$ to be an n-dimensional random vector whose components $\varepsilon_{t,i}$ have zero mean for each $i = 1,\ 2,\ \ldots,\ n$. Let $\mathbf{V}_\infty = \mathbb{E}(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t')$ be the n-by-n unconditional covariance matrix of $\boldsymbol{\varepsilon}_t$. Denote by $\boldsymbol{\Lambda}$ the diagonal matrix of the eigenvalues of $\mathbf{V}_\infty$,*

$$\boldsymbol{\Lambda} := \mathrm{diag}(\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_n),$$

ordered by descending values, $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$*. (The matrix* $\mathbf{V}_\infty$ *is positive definite, hence* $\lambda_i > 0$ *for any* i*.) Let* $\mathbf{P}$ *be the corresponding orthogonal matrix of normalized eigenvectors, so that the eigenvalue decomposition of* $\mathbf{V}_\infty$ *is given by*

$$\mathbf{V}_\infty = \mathbf{P}\boldsymbol{\Lambda}\mathbf{P}'.$$

Let further

$$\mathbf{L} := \mathbf{P}\boldsymbol{\Lambda}^{1/2}.$$

In other words, $\mathbf{L}$ *is an* n*-by-*n *matrix whose singular value decomposition is given by the product of an orthogonal matrix* $\mathbf{P}$*, a diagonal matrix* $\boldsymbol{\Lambda}^{1/2}$*, and the* n*-by-*n *identity matrix* $\mathbf{1}_n$*. Then,*

$$\mathbf{z}_t = \mathbf{L}^{-1}\boldsymbol{\varepsilon}_t, \tag{3.8}$$

is called the vector of principal components of $\boldsymbol{\varepsilon}_t$*, for any* t*. The* i*-th element of the vector* $\mathbf{z}_t$ *is called the* i-th principal component of $\boldsymbol{\varepsilon}_t$.

Note that

$$\mathbb{E}\left(\mathbf{z}_t\right) = \mathbf{L}^{-1}\mathbb{E}\left(\boldsymbol{\varepsilon}_t\right) = \mathbf{0}$$

and

$$\begin{aligned}
\operatorname{var}\left(\mathbf{z}_t\right) &= \mathbb{E}\left(\mathbf{z}_t\mathbf{z}_t'\right) \\
&= \mathbf{L}^{-1}\mathbb{E}\left(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\right)\mathbf{L}^{-1\prime} \\
&= \mathbf{L}^{-1}\mathbf{V}_\infty\mathbf{L}^{-1\prime} \\
&= \mathbf{1}_n, \qquad (3.9)
\end{aligned}$$

since $\mathbf{V}_\infty = \mathbf{L}\mathbf{L}'$. Hence, $\mathbf{z}_t$ are cross-sectionally uncorrelated and each component has a unit variance.

Since $\boldsymbol{\varepsilon}_t = \mathbf{L}\mathbf{z}_t$, each coordinate of $\boldsymbol{\varepsilon}_t$ can be written as a linear combination of the principal components,

$$\varepsilon_{t,i} = \sum_{j=1}^{n} L_{ij} z_{t,j}, \quad i = 1, 2, \ldots, n,$$

where L_{ij} are the elements of $\mathbf{L}$. The fraction of total variation in $\boldsymbol{\varepsilon}_t$ explained by the j-th principal component is

$$\frac{\lambda_j}{\sum_{k=1}^{n} \lambda_l}.$$

This property leads to another convenient feature of the principal component approach. Namely, if low-ranked components do not add much to the overall explained variance, which is often the case in financial time series, we can work with a reduced number of m principal components, where $m < n$. The first m components will then explain

$$\frac{\sum_{j=1}^{m} \lambda_j}{\sum_{k=1}^{n} \lambda_l} \lesssim 1$$

of the variation in $\boldsymbol{\varepsilon}_t$. In that case, $\mathbf{L}$ is replaced by a n-by-m matrix $\mathbf{L}_m$, where

$$\mathbf{L}_m := \mathbf{P}_m \boldsymbol{\Lambda}_m^{1/2}, \tag{3.10}$$

$\mathbf{P}_m$ is a n-by-m matrix of the first m normalized eigenvectors, and

$$\boldsymbol{\Lambda} := \mathrm{diag}(\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_m)$$

is a diagonal matrix of the first m eigenvalues. The m-dimensional vector of the first m principal components of $\boldsymbol{\varepsilon}_t$ is then given by

$$\mathbf{z}_t = \mathbf{L}_m^{-1} \boldsymbol{\varepsilon}_t, \tag{3.11}$$

for any t.

Filtering

Orthogonalization transforms a cross-sectionally correlated series into a set of uncorrelated ones. We also have to filter out any serial correlation and volatility clustering. As a net result we will obtain sequences of orthogonal, serially uncorrelated and identically distributed conditional residuals.

Specifically, I will assume that for each asset $i = 1, 2, \ldots, n$ the log returns $y_{t,i} := \ln(S_{t,i}/S_{t-1,i})$ at time t follow an ARMA(r,m) process

$$y_{t,i} = \mu_i + \sum_{s=1}^{r} b_{s,i} y_{t-s,i} + \varepsilon_{t,i} + \sum_{s=1}^{m} \theta_{s,i} \varepsilon_{t-s,i}. \tag{3.12}$$

For each t and i, the residuals $\varepsilon_{t,i}$ are serially uncorrelated random variables with a continuous density function of zero mean. Conditionally on the information available at $t-1$, the vector of residuals,

$$\boldsymbol{\varepsilon}_t := [\varepsilon_{t,1} \quad \varepsilon_{t,2} \quad \ldots \quad \varepsilon_{t,n}]',$$

has a zero mean and a covariance matrix $\mathbf{V}_t$. That is,

$$\begin{aligned} \mathbb{E}\left(\boldsymbol{\varepsilon}_t | \mathcal{F}_{t-1}\right) &= \mathbb{E}\left(\boldsymbol{\varepsilon}_t\right) = [0 \quad 0 \quad \ldots \quad 0]' =: \mathbf{0}, && (3.13) \\ \operatorname{var}\left(\boldsymbol{\varepsilon}_t | \mathcal{F}_{t-1}\right) &= \mathbb{E}\left(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' | \mathcal{F}_{t-1}\right) =: \mathbf{V}_t, && (3.14) \end{aligned}$$

where, for any t, the matrix $\mathbf{V}_t$ is positive definite and measurable with respect to the information set $\mathcal{F}_{t-1}$, a σ-field generated by the past residuals $\{\boldsymbol{\varepsilon}_{t-1}, \ \boldsymbol{\varepsilon}_{t-2}, \ \ldots, \ \boldsymbol{\varepsilon}_1\}$. Note that the vector form of the ARMA process given by equation (3.12) then reads

$$\mathbf{y}_t = \boldsymbol{\mu} + \sum_{s=1}^{r} \mathbf{b}_s \mathbf{y}_{t-s} + \boldsymbol{\varepsilon}_t + \sum_{s=1}^{m} \boldsymbol{\theta}_s \boldsymbol{\varepsilon}_{t-s}, \tag{3.15}$$

where $\mathbf{y}_t$ and $\boldsymbol{\mu}$ are vectors with elements indexed by $i = 1, 2, \ldots, n$, while

$$\mathbf{b}_s \ := \ \operatorname{diag}\left(b_{s,1}, \ b_{s,2}, \ \ldots, \ b_{s,n}\right)$$

$$\boldsymbol{\theta}_s \;:=\; \operatorname{diag}\left(\theta_{s,1},\ \theta_{s,2},\ \ldots,\ \theta_{s,n}\right)$$

are n-by-n diagonal matrices of ARMA coefficients.

To capture the volatility clustering, I will assume that the conditional covariance matrix follows a model from the GARCH family. The standard GARCH(p,q) model is sufficient to capture most of the clustering, and – to some extent – excess kurtosis. However, it has a drawback of being symmetric, in the sense that negative and positive shocks have the same impact on volatility. There is a strong empirical evidence that the positive and negative innovations to returns exhibit different correlations with innovations to volatility. This asymmetry can be captured, for example, by assuming that the conditional residuals follow an asymmetric distribution, such as skewed Student's t. Alternatively, we can model the asymmetry explicitly in the equation followed by the conditional covariance matrix. In order to keep the estimation method free of any distributional assumptions I opt for the alternative approach. As Glosten, Jagannathan, and Runkle (1993), I will assume that the conditional covariance $\mathbf{V}_t$ follows a multivariate asymmetric GARCH(p,q), also known as multivariate GJR-GARCH(p,q):

$$\mathbf{V}_t = \boldsymbol{\Omega} + \sum_{s=1}^{p} \mathbf{A}_s \mathbf{E}_{t-s} + \sum_{s=1}^{p} \boldsymbol{\Theta}_s \mathbf{I}_{t-s} \mathbf{E}_{t-s} + \sum_{s=1}^{q} \mathbf{B}_s \mathbf{V}_{t-s}, \tag{3.16}$$

where $\boldsymbol{\Omega}$, $\mathbf{A}_1$, $\ldots$, $\mathbf{A}_p$, $\boldsymbol{\Theta}_1$, $\ldots$, $\boldsymbol{\Theta}_p$, $\mathbf{B}_1$, $\ldots$, $\mathbf{B}_q$ are constant, positive semidefinite n-by-n matrices,

$$\mathbf{E}_t := \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t',$$

and

$$\mathbf{I}_t := \operatorname{diag}(\operatorname{sgn}(-\varepsilon_{t,1})_+,\ \operatorname{sgn}(-\varepsilon_{t,2})_+,\ \ldots,\ \operatorname{sgn}(-\varepsilon_{t,n})_+),$$

for any t. As usual, the coefficients in matrices $\mathbf{A}_s$ in (3.16) measure the extent to which volatility shocks in previous periods affect the current volatility, while $\mathbf{A}_s + \mathbf{B}_s$ measure the rate at which this effect fades away. The terms proportional to matrices $\boldsymbol{\Theta}_s$ capture the impact of asymmetric return shocks to volatility. For any t, the unconditional covariance matrix of $\boldsymbol{\varepsilon}_t$ is given by

$$\mathbf{V}_\infty := \left(\mathbf{1}_n - \sum_{s=1}^{p}\left(\mathbf{A}_s + \frac{1}{2}\boldsymbol{\Theta}_s\right) - \sum_{s=1}^{q}\mathbf{B}_s\right)^{-1}\boldsymbol{\Omega}.$$

Hence, covariance stationarity of the GJR-GARCH(p,q) process (3.16) is assured by setting the matrix

$$\mathbf{1}_n - \sum_{s=1}^{p}\left(\mathbf{A}_s + \frac{1}{2}\boldsymbol{\Theta}_s\right) - \sum_{s=1}^{q}\mathbf{B}_s$$

to be positive definite.

It is worth noting that there are many plausible and often implemented alternatives to asymmetric GARCH model of Glosten, Jagannathan, and Runkle (1993), such as EGARCH model of Nelson (1991). I have chosen to work with the Glosten, Jagannathan, and Runkle (1993) specification for the sake of simplicity.

Cross-sectional correlations are reflected in the off-diagonal terms of matrices $\mathbf{V}_t$ and $\mathbf{E}_t$. This in turn makes the matrices $\boldsymbol{\Omega}$, $\mathbf{A}_1, \ \ldots, \ \mathbf{A}_p$, $\boldsymbol{\Theta}_1, \ \ldots, \ \boldsymbol{\Theta}_p$, $\mathbf{B}_1, \ \ldots, \ \mathbf{B}_q$ non-diagonal. In total, one would have to estimate $(1+2p+q)(n+1)n/2$ different parameters. Clearly, this number explodes as we increase the number of assets in the portfolio. However, this is only one facet of the problem. The other is that we cannot apply the results of univariate EVT to conditional residuals $\varepsilon_{t,i}$ directly.

For that matter, we can work in the orthonormal basis of principal components by applying the linear transformation (3.8) to the conditional residuals $\boldsymbol{\varepsilon}_t$. In the orthonormal basis of principal components, equation (3.15) and (3.16) for the ARMA(r,m)–GJR-GARCH(p,q) process then read:

$$\widehat{\mathbf{y}}_t = \widehat{\boldsymbol{\mu}} + \sum_{s=1}^{r}\widehat{\mathbf{b}}_s\widehat{\mathbf{y}}_{t-s} + \mathbf{z}_t + \sum_{s=1}^{m}\widehat{\boldsymbol{\theta}}_s\mathbf{z}_{t-s} \tag{3.17}$$

and

$$\widehat{\mathbf{V}}_t = \widehat{\mathbf{\Omega}} + \sum_{s=1}^{p} \widehat{\mathbf{A}}_s \widehat{\mathbf{E}}_{t-s} + \sum_{s=1}^{p} \widehat{\mathbf{\Theta}}_s \widehat{\mathbf{I}}_{t-s} \widehat{\mathbf{E}}_{t-s} + \sum_{s=1}^{q} \widehat{\mathbf{B}}_s \widehat{\mathbf{V}}_{t-s}, \tag{3.18}$$

where $\widehat{\mathbf{y}}_t := \mathbf{L}^{-1}\mathbf{y}_t$ for any t, $\widehat{\boldsymbol{\mu}} := \mathbf{L}^{-1}\boldsymbol{\mu}$, and

$$\widehat{\mathbf{M}} := \mathbf{L}^{-1}\mathbf{M}\mathbf{L}^{-1\prime}$$

for any $\mathbf{M} \in \{\widehat{\mathbf{b}};\ \mathbf{\Omega},\ \mathbf{A}_1,\ \ldots,\ \mathbf{A}_p,\ \mathbf{\Theta}_1,\ \ldots,\ \mathbf{\Theta}_p,\ \mathbf{B}_1,\ \ldots,\ \mathbf{B}_q\}$ and any $\mathbf{M} \in \{\boldsymbol{V}_t,\ \mathbf{E}_t,\ \mathbf{I}_t\}_{t \geq \max\{p,q\}}$. In particular,

$$\widehat{\mathbf{E}}_t := \mathbf{L}^{-1}\mathbf{E}_t\mathbf{L}^{-1\prime} = \mathbf{z}_t\mathbf{z}_t'$$

and

$$\widehat{\mathbf{I}}_t := \mathbf{L}^{-1}\mathbf{I}_t\mathbf{L}^{-1\prime} = \operatorname{diag}(\operatorname{sgn}(-z_{t,1})_+,\ \operatorname{sgn}(-z_{t,2})_+,\ \ldots,\ \operatorname{sgn}(-z_{t,n})_+).$$

Equation (3.13) implies

$$\mathbb{E}\left(\mathbf{z}_t | \mathcal{F}_{t-1}\right) = \mathbf{L}^{-1}\mathbb{E}\left(\boldsymbol{\varepsilon}_t\right) = \mathbf{0}. \tag{3.19}$$

On the other hand, let

$$\begin{aligned} \widehat{\mathbf{V}}_t &:= \operatorname{var}\left(\mathbf{z}_t | \mathcal{F}_{t-1}\right) \\ &= \mathbb{E}\left(\mathbf{z}_t\mathbf{z}_t' | \mathcal{F}_{t-1}\right) \\ &= \mathbf{L}^{-1}\mathbf{V}_t\mathbf{L}^{-1\prime} \end{aligned}$$

be the conditional covariance matrix of principal components. Since the principal components $\mathbf{z}_t$ are orthogonal, it is reasonable to assume that the matrix $\widehat{\mathbf{V}}_t$ is diagonal (see, for example, Alexander (2001)). Then, the process given by equation (3.18) can be estimated separately for each principal component.

This gives a set of n independent scalar equations of the form

$$\widehat{V}_{t,i} = \widehat{\Omega}_i + \sum_{s=1}^{p} \widehat{A}_{s,i}\widehat{E}_{t-s,i} + \sum_{s=1}^{p} \widehat{\Theta}_{s,i}\widehat{I}_{t-s,i}\widehat{E}_{t-s,i} + \sum_{s=1}^{q} \widehat{B}_{s,i}\widehat{V}_{t-s,i}, \tag{3.20}$$

where, in general, $\widehat{M}_i := \widehat{M}_{ii}$ is the i-th diagonal element of the matrix $\widehat{\mathbf{M}}$, i being $1, 2, \ldots, n$ for the first, second, …, n-th principal component, respectively.

Once we estimate the set of parameters $\{\widehat{\mathbf{\Omega}},\ \widehat{\mathbf{A}}_1,\ \ldots,\ \widehat{\mathbf{A}}_p,\ \widehat{\mathbf{\Theta}}_1,\ \ldots,\ \widehat{\mathbf{\Theta}}_p,\ \widehat{\mathbf{B}}_1,\ \ldots,\ \widehat{\mathbf{B}}_q\}$ we can apply the inverse transformation

$$\mathbf{V}_t := \mathbf{L}\widehat{\mathbf{V}}_t\mathbf{L}' \tag{3.21}$$

for $t \geq \max\{p, q\}$, to retrieve the series of conditional covariance matrices in the original basis of log returns. This allows us to estimate VaR and ES in a multivariate framework, for an arbitrary portfolio.

Note that it is straightforward to generalize the above approach to the case of $m < n$ principal components. Using definition (3.10), we can transform any n-by-n matrix $\mathbf{M}$ into the basis of the first m principal components via transformation

$$\widehat{\mathbf{M}} := \mathbf{L}_m^{-1}\mathbf{M}\mathbf{L}_m^{-1\prime},$$

yielding an m-by-m matrix $\widehat{\mathbf{M}}$. Equations (3.18) and (3.20) maintain the same form.

GMM Estimation

Estimation of the GJR-GARCH(p, q) parameters in the basis of principal components can be performed in several ways. Let us focus on the set of scalar equations (3.20). Under the additional assumption of a known conditional distribution for the residuals, it is straightforward to set up the likelihood function for the entire ARMA(r, m)–GJR-GARCH(p, q) model. This gives the

ML estimator for the set of parameters

$$\left\{\widehat{\boldsymbol{\mu}},\ \widehat{\mathbf{b}}_1,\ \ldots,\ \widehat{\mathbf{b}}_r,\ \widehat{\boldsymbol{\theta}}_1,\ \ldots,\ \widehat{\boldsymbol{\theta}}_m;\ \widehat{\boldsymbol{\Omega}},\ \widehat{\mathbf{A}}_1,\ \ldots,\ \widehat{\mathbf{A}}_p,\ \widehat{\boldsymbol{\Theta}}_1,\ \ldots,\ \widehat{\boldsymbol{\Theta}}_p,\ \widehat{\mathbf{B}}_1,\ \ldots,\ \widehat{\mathbf{B}}_q\right\}.$$

In the principal component framework, there are $(1 + r + m + 1 + 2p + q)n$ parameters in total to be estimated.

However, as indicated earlier, it is desirable to have an estimator which avoids specific assumptions about the conditional distribution, while maintaining the efficiency of the ML (or quasi-ML) estimator. Such an estimator is based on the Generalized Method of Moments (GMM). Instead of making distributional assumptions, it proceeds by postulating conditional moments. Here, I will briefly outline its implementation. The details of the GMM approach to ARMA-GARCH models can be found, for example, in Skoglund (2001).

For a fixed t and any principal component i define

$$\mathbf{e}_t := \left[z_t \quad z_t^2 - \widehat{V}_t\right]',$$

where, with a slight abuse of notation, I use $z_t = z_{t,i}$ and $\widehat{V}_t = \widehat{V}_{t,i}$. Let the score be given by

$$\mathbf{g}_t(\psi) := \mathbf{F}_t'(\psi)\mathbf{e}_t,$$

where $\mathbf{F}_t$ is an instrumental variable function. The GMM estimator of univariate ARMA(r, m)–GJR-GARCH(p, q) parameters ψ is defined as

$$\widehat{\psi} = \arg\min_{\psi}\ \mathbf{m}(\psi)'\ \mathbf{W}\ \mathbf{m}(\psi), \tag{3.22}$$

where

$$\mathbf{m}(\psi) := \frac{1}{T}\sum_{t=1}^{T} \mathbf{g}_t(\psi)$$

is the sample analog of the expected score, while the weighting matrix $\mathbf{W}$ is a consistent estimate of the inverse asymptotic covariance matrix of the score.

The set of moment conditions is given by

$$\mathbb{E}\left[\mathbf{g}_t(\psi)\right] = \mathbf{0}.$$

An efficient choice of instrumental variable function and weighting matrix corresponds to setting

$$\mathbf{F}_t(\psi) = \mathbf{\Sigma}_t^{-1}\mathbf{J}_t(\psi),$$

where $\mathbf{\Sigma}_t := \mathrm{var}\left(\mathbf{e}_t|\mathcal{F}_{t-1}\right)$,

$$\mathbf{J}_t(\psi) := \frac{\partial \mathbf{e}_t}{\partial \psi'}$$

is the Jacobian matrix, and

$$\mathbf{W} = \frac{1}{T}\sum_{t=1}^{T}\mathbf{g}_t(\psi)\mathbf{g}_t(\psi)',$$

see Newey and McFadden (1994). Standard errors can be calculated in a usual way from a consistent estimate of the Fisher information matrix. A recursive semi-closed form solution for $\mathbf{g}_t$ can be found in Skoglund (2001), Nyström and Skoglund (2002b) and Nyström and Skoglund (2002b), for a particular (and common) case of ARMA$(1,0)$–GJR-GARCH$(1,1)$ process.

It is worth noting that the application of the GMM estimator requires an initial guess on the third and fourth moments of z_t. Therefore, in order to obtain an initial estimator of the set of parameters ψ, it is convenient to use the quasi-ML estimator to obtain the initial consistent estimates (i.e., to assume conditional normality of residuals).

Forecasting

A one-step-ahead forecast of the transformed log return vector can be obtained from (3.17):

$$\mathbb{E}\left(\widehat{\mathbf{y}}_{t+1}|\mathcal{F}_t\right) = \widehat{\boldsymbol{\mu}} + \sum_{s=1}^{r} \widehat{\mathbf{b}}_s \widehat{\mathbf{y}}_{t-s+1}. \tag{3.23}$$

Using the fact that principal components $\mathbf{z}_t$ are independent, we can write the forecast for an arbitrary time horizon $h \geq 1$:

$$\begin{aligned}\mathbb{E}\left(\widehat{\mathbf{y}}_{t+h}|\mathcal{F}_t\right) &= \widehat{\boldsymbol{\mu}} + \sum_{s=1}^{r} \widehat{\mathbf{b}}_s \mathbb{E}\left(\widehat{\mathbf{y}}_{t-s+h}|\mathcal{F}_t\right) \\ &= \widehat{\boldsymbol{\mu}} + \sum_{s=1}^{h} \widehat{\mathbf{b}}_s \widehat{\mathbf{y}}_{t-s+h} + \sum_{s=h+1}^{r} \widehat{\mathbf{b}}_s \mathbb{E}\left(\widehat{\mathbf{y}}_{t-s+h}|\mathcal{F}_t\right). \end{aligned} \tag{3.24}$$

Equation (3.24) is recursive and the last term contains the forecasts for $1,\ 2,\ \ldots,\ h-1$ steps ahead.

Next, from equation (3.18), it follows that a one-step-ahead forecast of conditional covariance in the basis of principal components is given by

$$\begin{aligned}\mathbb{E}\left(\widehat{\mathbf{V}}_{t+1}|\mathcal{F}_t\right) &= \widehat{\mathbf{V}}_{t+1} \\ &= \widehat{\boldsymbol{\Omega}} + \sum_{s=1}^{p} \widehat{\mathbf{A}}_s \widehat{\mathbf{E}}_{t-s+1} + \sum_{s=1}^{p} \widehat{\boldsymbol{\Theta}}_s \widehat{\mathbf{I}}_{t-s+1} \widehat{\mathbf{E}}_{t-s+1} + \sum_{s=1}^{q} \widehat{\mathbf{B}}_s \widehat{\mathbf{V}}_{t-s+1},\end{aligned}$$

since $\widehat{\mathbf{V}}_{t+1}$ is measurable with respect to the information available at t. A two-steps-ahead forecast is

$$\mathbb{E}\left(\widehat{\mathbf{V}}_{t+2}|\mathcal{F}_t\right) = \widehat{\boldsymbol{\Omega}} + \sum_{s=1}^{p} \left(\widehat{\mathbf{A}}_s + \frac{1}{2}\widehat{\boldsymbol{\Theta}}_s\right) \widehat{\mathbf{V}}_{t-s+1} + \sum_{s=1}^{q} \widehat{\mathbf{B}}_s \widehat{\mathbf{V}}_{t-s+1},$$

which can be obtained by substituting the matrices known up until and including time t. Iteratively, we can derive a covariance forecast for an arbitrary horizon. Applying the inverse transformation (3.21), we can obtain the covariance forecast in the original basis of log returns.

Given the upper- and lower-tail quantiles $q_{\pm}$, the confidence interval $[z^{-}_{t+h,i},\ z^{+}_{t+h,i}]$ for the forecast of the value of i-th principal component h steps ahead is given by

$$z^{\pm}_{t+h,i} = F_i^{-1}(q_{\pm})\sqrt{\widehat{V}_{t+h,i}}, \tag{3.25}$$

where, $F_i^{-1}(\cdot)$ is the inverse of the univariate probability function followed by the set of random variables $\{z_{t,i}\}_{t=1}^{T}$. It can be obtained by inverting the tail estimator (3.3). As before, $\widehat{V}_{t+h,i}$ stands for the i-th diagonal element of the matrix $\widehat{\mathbf{V}}_{t+h}$.

Estimating Multivariate VaR and ES

Our final goal is to estimate VaR and ES for a portfolio of n assets. Denote by $\mathbf{a}$ the vector of portfolio positions, in monetary units.[2] Then, h-steps-ahead portfolio VaR is defined by

$$\mathrm{VaR}_{q_{\pm}} = \mathbf{a}'\mathbf{L}\left[\mathbb{E}\left(\widehat{\mathbf{y}}_{t+h}|\mathcal{F}_t\right) + \mathbf{z}^{\pm}_{t+h}\right], \tag{3.26}$$

where $\mathbf{z}^{\pm}_{t+h}$ is the vector whose i-th component is given by (3.25). The intuition behind formula (3.26) is the following. The first term,

$$\mathbf{a}'\mathbf{L}\ \mathbb{E}\left(\widehat{\mathbf{y}}_{t+h}|\mathcal{F}_t\right) = \mathbf{a}'\mathbb{E}\left(\mathbf{y}_{t+h}|\mathcal{F}_t\right),$$

represents the expected return on the portfolio for h steps ahead. The second term is determined by the vector $\mathbf{L}\mathbf{z}^{\pm}_{t+h}$, which defines the confidence intervals in the n-dimensional space of log returns. Hence, the second term $\mathbf{a}'\mathbf{L}\mathbf{z}^{\pm}_{t+h}$ is the confidence interval for portfolio returns around their mean, for h steps ahead and at a confidence level defined by $q_{\pm}$.

In analogy with equations (3.26) and (3.5), the h-steps-ahead portfolio ES

[2] This, among other things, facilitates the treatment of short positions, when portfolio weights may not be well defined.

is given by

$$\mathrm{ES}_{q\pm} = \mathbf{a}'\mathbf{L}\widetilde{\mathbf{z}}^{\pm}_{t+h}, \tag{3.27}$$

where

$$\widetilde{z}^{\pm}_{t+h,i} = \widetilde{F}_i^{-1}(q_\pm)\sqrt{\widehat{V}_{t+h,i}}$$

and

$$\widetilde{F}_i^{-1}(q_\pm) = \frac{1}{1-\xi_\pm}\left[F_i^{-1}(q_\pm) \pm \beta_\pm - \xi_\pm u_\pm\right].$$

3.3 Data and Empirical Results

3.3.1 Data

The empirical results that follow are based on average daily interbank spot exchange rates of Euro, British Pound, Japanese Yen and Swiss Franc with respect to the U.S. Dollar, from January 4, 1999 to September 30, 2008, a sample of 2542 observations. The four time series were obtained from Thomson Financial's Datastream. Table 3.1 provides summary statistics for the exchange rate levels and the corresponding daily log returns (in percent), computed as $y_t = 100\ln(S_t/S_{t-1})$. Daily sampling is chosen in order to capture high-frequency fluctuations in return processes that may be critical for identification of rare events in the tails of distribution, while avoiding to model the intraday return dynamics, abundant with spurious market microstructure distortions and trading frictions.

I perform several preliminary test on the data. The values of skewness and kurtosis in Table 3.1 indicate that both the levels and returns deviate from normality. This is also confirmed by Jarque-Bera and Kolmogorov-Smirnov tests (not reported), whose p-values are at most of the order of 10^{-3}. Table 3.2 shows the results of Ljung-Box Q-statistics for the autocorrelation of returns, up to order 10 (Panel A). The null hypotheses of no autocorrelation in returns cannot be rejected. The absence of a significant short-run return predictability

Table 3.1: **Summary Statistics**

Daily interbank spot exchange rates of Euro, British Pound, Japanese Yen and Swiss Franc with respect to the U.S. Dollar, from January 4, 1999 to September 30, 2008 (2542 observations).

Panel A: Daily exchange rate levels				
Currency	Mean	Variance	Skewness	Kurtosis
EUR	1.1511	0.0376	0.2234	2.1992
GBP	1.7103	0.0376	0.0661	1.7385
JPY	0.8774	0.0030	−0.0339	2.2511
CHF	0.7380	0.0121	0.1375	2.1955

Panel B: Daily returns (percent)				
Currency	Mean	Variance	Skewness	Kurtosis
EUR	0.0084	0.3539	−0.0267	4.5420
GBP	0.0040	0.2338	0.0757	4.1778
JPY	0.0026	0.3493	0.2267	4.8656
CHF	0.0088	0.4012	0.1411	4.2532

is consistent with high efficiency of the currency market. The autocorrelation in the squared returns is, on the other hand, highly significant in all four series, indicating the presence of heteroskedasticity (Panel B). The correlation coefficients between squared returns and their lags (not reported) are all positive, confirming the notion of clustering – the periods of high volatility are likely to be followed by high volatility.

Table 3.3 reports the results of the unit root tests. Both Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) statistics indicate that the unit root hypothesis is convincingly rejected in favor of stationary returns (the critical values of ADF and PP statistics at 5 and 1 percent confidence are -3.41 and -3.96, respectively).

Table 3.2: **Autocorrelation**

Ljung-Box test for autocorrelation of returns and squared returns up to 10^{th} lag.

Panel A: Autocorrelation of returns		
Currency	Q statistic	p-value
EUR	3.9867	0.9479
GBP	9.4858	0.4867
JPY	6.8611	0.7385
CHF	12.7326	0.2390

Panel B: Autocorrelation of squared returns		
Currency	Q statistic	p-value
EUR	111.5435	$< 10^{-5}$
GBP	105.7946	$< 10^{-5}$
JPY	81.5108	$< 10^{-5}$
CHF	42.7107	$< 10^{-5}$

Table 3.3: **Stationarity**

Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests for the presence of unit roots, based on the regression

$$y_t = c + \delta t + \phi y_{t-1} + \sum_{L=1}^{10} b_L \Delta y_{t-L} + \varepsilon_t,$$
$$\mathrm{H}_0 : \phi = 1,\ \delta = 0.$$

Currency	ADF statistic	PP statistic
EUR	−15.8257	−50.4768
GBP	−15.2648	−48.0508
JPY	−14.6802	−49.2301
CHF	−15.6345	−50.9751
5% crit. value	−3.41	−3.41
1% crit. value	−3.96	−3.96

3.3.2 Empirical Results

I apply the method described in Section 3.2 to the exchange rate data. The ex-post analysis of autocorrelations in principal components and squared principal components have shown that it is sufficient to use an ARMA$(1,0)$–GJR-GARCH$(1,1)$ model to obtain independent and identically distributed residuals. Hence, the estimation steps are the following. First, estimate

$$\mathbf{y}_t = \boldsymbol{\mu} + \mathbf{b}\mathbf{y}_{t-s} + \boldsymbol{\varepsilon}_t. \tag{3.28}$$

Then, calculate the unconditional variance matrix $\mathbf{V}_\infty$ of the residuals $\boldsymbol{\varepsilon}_t$ and apply the eigenvalue decomposition following Definition 2 to obtain the principal components $\mathbf{z}_t$. The conditional covariance matrix of the principal components then follows

$$\widehat{\mathbf{V}}_t = \widehat{\boldsymbol{\Omega}} + \widehat{\mathbf{A}}\widehat{\mathbf{E}}_{t-1} + \widehat{\boldsymbol{\Theta}}\widehat{\mathbf{I}}_{t-1}\widehat{\mathbf{E}}_{t-1} + \widehat{\mathbf{B}}\widehat{\mathbf{V}}_{t-1}. \tag{3.29}$$

Next, to obtain the GJR-GARCH$(1,1)$ parameters

$$\boldsymbol{\psi} := \left\{\widehat{\boldsymbol{\Omega}},\ \widehat{\mathbf{A}},\ \widehat{\boldsymbol{\Theta}},\ \widehat{\mathbf{B}}\right\},$$

run the GMM estimation (3.22) separately for each principal component. Since there are four exchange rates in the sample, there are $4 \times 4 = 16$ parameters in total to be estimated from the GMM step if we work with a full set of four principal components.

Once we have the ARMA$(1,0)$–GJR-GARCH$(1,1)$ parameters, we can compute the forecasts, as well as VaR and ES for an arbitrary portfolio, following formulas (3.26) and (3.27). Note that, in general, parameter estimates change in time as we move through the time series. Hence, a proper dynamic method for VaR and ES forecasting would involve regular updating of parameters.

I illustrate the method by running a dynamic estimation over the sample period. To have sufficient observations for the estimation runs, I start from

January 1, 2004. For each of the remaining 1239 daily observations, I calculate one-step-ahead forecasts of VaR and ES. As an example, the estimation details are shown in Tables 3.4–3.7 and Figures 3.1–3.4, for January 1, 2008.

Table 3.4 shows the summary of the principal component analysis. The first principal component explains almost 70 percent of joint variations in the four exchange rates.

Table 3.4: **Principal Components**

Variance explained by each of the four principal components (PCs). Estimation period is January 4, 1999 – December 31, 2007.

	PC 1	PC 2	PC 3	PC 4
Eigenvalue	0.9283	0.2682	0.1036	0.0269
Variance explained	0.6995	0.2021	0.0781	0.0202
Cumulative	0.6995	0.9017	0.9798	1.0000

Table 3.5 summarizes the results of the univariate parameter estimation. For each principal component i, the table shows the values of ARMA(1,0)–GJR-GARCH(1,1) parameters obtained by the GMM estimation, along with their standard errors. Clearly, mean stationarity is satisfied, since $|\widehat{b}_i| < 1$ for every i. Also, it is easy to check that the GARCH parameters are very close but still within the bounds of covariance-stationary regime, as $\widehat{A}_i + \widehat{\Theta}_i/2 + \widehat{B}_i < 1$. The constant terms $\widehat{\Omega}_i$ appear to be insignificant; however, the corresponding unconditional variances $\widehat{V}_{\infty,i} = \widehat{\Omega}_i/[1 - (\widehat{A}_i + \widehat{\Theta}_i/2 + \widehat{B}_i)]$ are significant.

Table 3.6 displays the estimates of the upper- and lower-tail parameters of the univariate GP distribution, $\xi_\pm$ and $\beta_\pm$. Following the procedure described in Subsection 3.2.2, these parameters are estimated separately for each of the principal components, as the exceedances can be assumed to be not only independent and identically distributed, but also orthogonal. The upper and lower tresholds, u_+ and u_-, are determined by $F(u_+) = 0.90$ and $F(u_-) = 0.10$, respectively. This gives a sufficient number of observations in the tails to render

Table 3.5: **ARMA-GARCH Estimates**

Parameter estimates in the ARMA(1,0)–GJR-GARCH(1,1) model, in the basis of principal components. Estimation period is January 4, 1999 – December 31, 2007.

Parameter	PC 1	PC 2	PC 3	PC 4
$\widehat{\mu}_i$	0.0018	0.0043	0.0034	−0.0002
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\widehat{b}_i$	0.0212	0.0777	0.0910	0.0093
	(0.0077)	(0.0112)	(0.0136)	(0.0443)
$\widehat{\Omega}_i$	0.0005	0.0039	0.0016	0.0016
	(0.0092)	(0.0215)	(0.0179)	(0.0407)
$\widehat{A}_i$	0.0216	0.0446	0.0361	0.1593
	(0.0069)	(0.0238)	(0.0210)	(0.0594)
$\widehat{\Theta}_i$	0.0030	0.0162	0.0007	−0.0942
	(0.0002)	(0.0001)	(0.0001)	(0.0001)
$\widehat{B}_i$	0.9751	0.9218	0.9350	0.6337
	(0.0201)	(0.0185)	(0.0196)	(0.0139)

(Standard errors in parentheses.)

the ML estimation of the parameters possible. The inverse of the tail index, $\xi_{\pm}$, is significant and negative for the first principal component, which corresponds to Weibull distribution. The other values of $\xi_{\pm}$ are statistically not different from zero, with an exception of the upper tail of the fourth principal component, where ξ_{+} is significant and positive. The scale parameter $\beta_{\pm}$ has values that range between 0.52 and 0.70. Also, the asymmetry between the upper and the lower tail implied by the parameters is apparent.

Using formula (3.1) and the values in Table 3.6, we can plot the function $G_{\xi,\beta}(\cdot)$ for the distribution of excesses of $z_{t,i}$ over the upper and lower tresholds, u_{+} and u_{-}. Figures 3.1–3.4 show the graphs for the tail behavior of each of the four principal components. I compare the empirical with the GP distribution function, as well as with the normal and Student's t distributions calibrated across the sample, using the parameter estimates prior to January 1, 2008. Clearly, GP distribution drastically outperforms the alternatives in explaining the tail behavior.

Table 3.6: **Parameters of the Univariate GP Distribution**

Upper- and lower-tail parameters of the univariate generalized Pareto distribution, estimated separately for each of the standardized ARMA-GARCH orthogonal residuals. The upper and lower tresholds are determined by the quantiles corresponding to probabilities of 0.90 and 0.10, respectively. Estimation period is January 4, 1999 – December 31, 2007.

Upper tail				
Parameter	PC 1	PC 2	PC 3	PC 4
$\widehat{\xi}_+$	−0.1096	0.0544	0.0058	0.1804
	(0.0612)	(0.0703)	(0.0755)	(0.0752)
$\widehat{\beta}_+$	0.6397	0.5153	0.5724	0.5386
	(0.0573)	(0.0497)	(0.0575)	(0.0536)

Lower tail				
Parameter	PC 1	PC 2	PC 3	PC 4
$\widehat{\xi}_-$	−0.2030	0.0570	−0.0293	0.0239
	(0.0575)	(0.0714)	(0.0572)	(0.0625)
$\widehat{\beta}_-$	0.7013	0.6765	0.6379	0.6031
	(0.0605)	(0.0658)	(0.0556)	(0.0548)

(Standard errors in parentheses.)

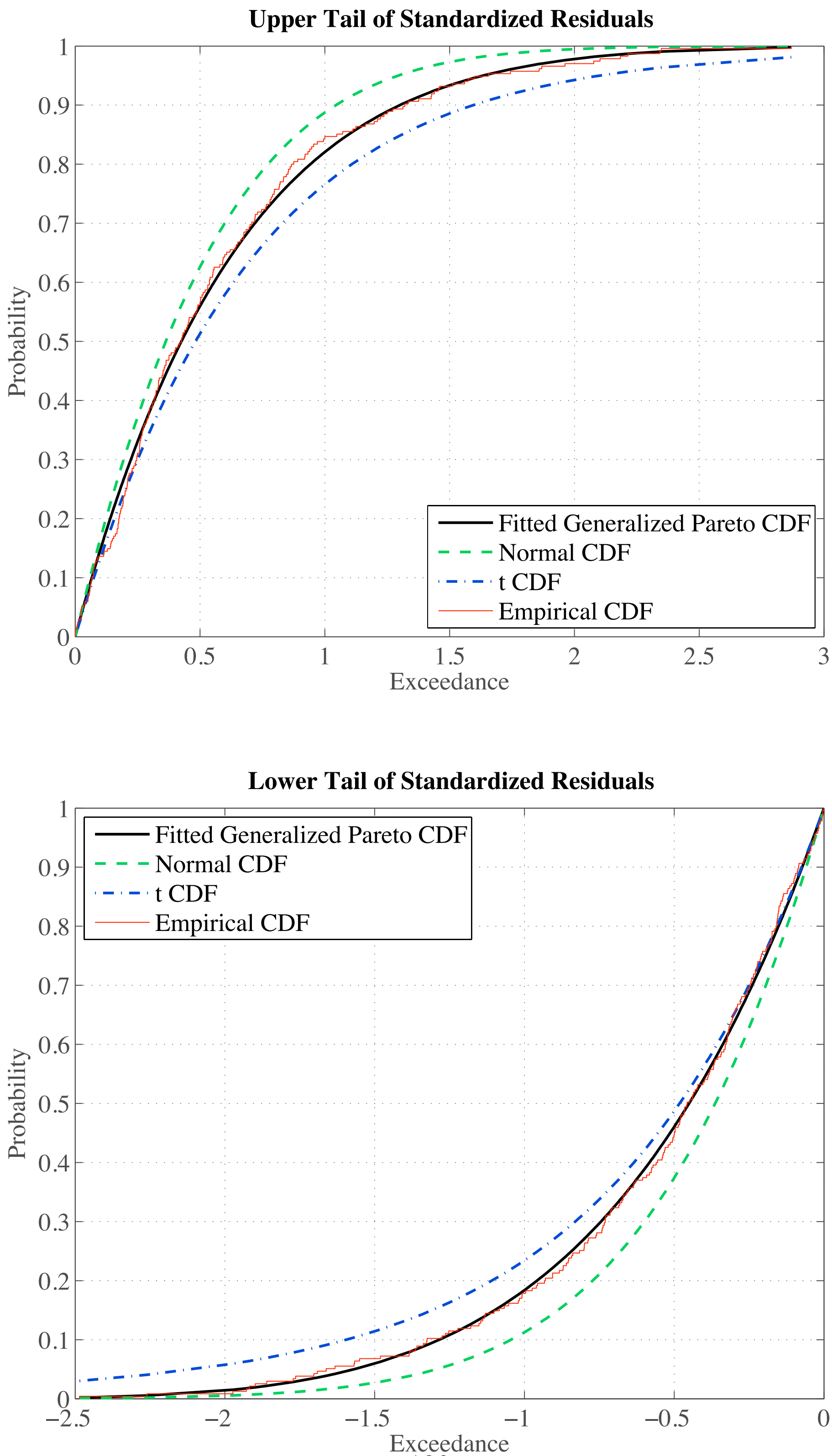

Figure 3.1: **First principal component.** Upper and lower tails of standardized residuals.

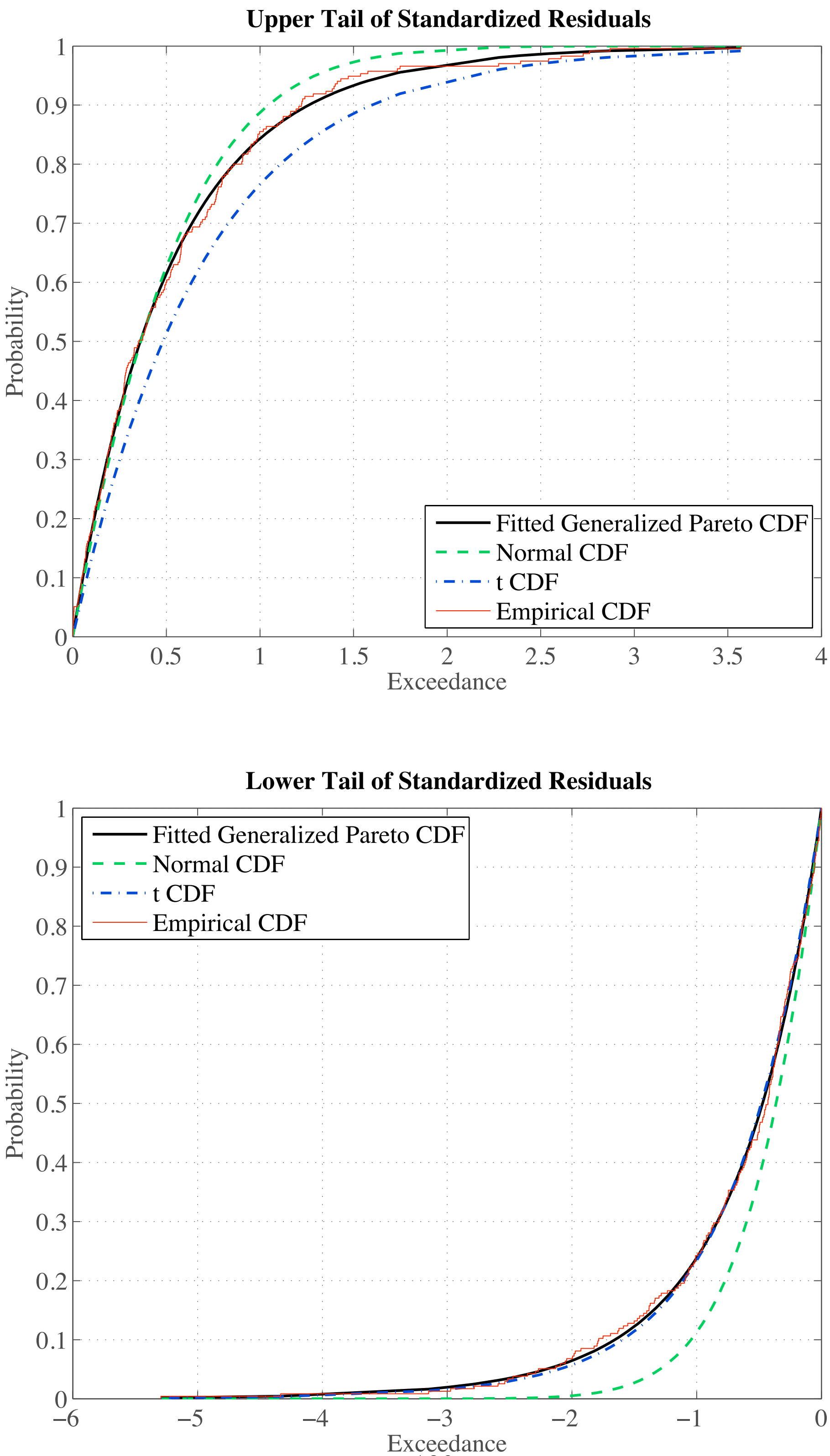

Figure 3.2: **Second principal component.** Upper and lower tails of standardized residuals.

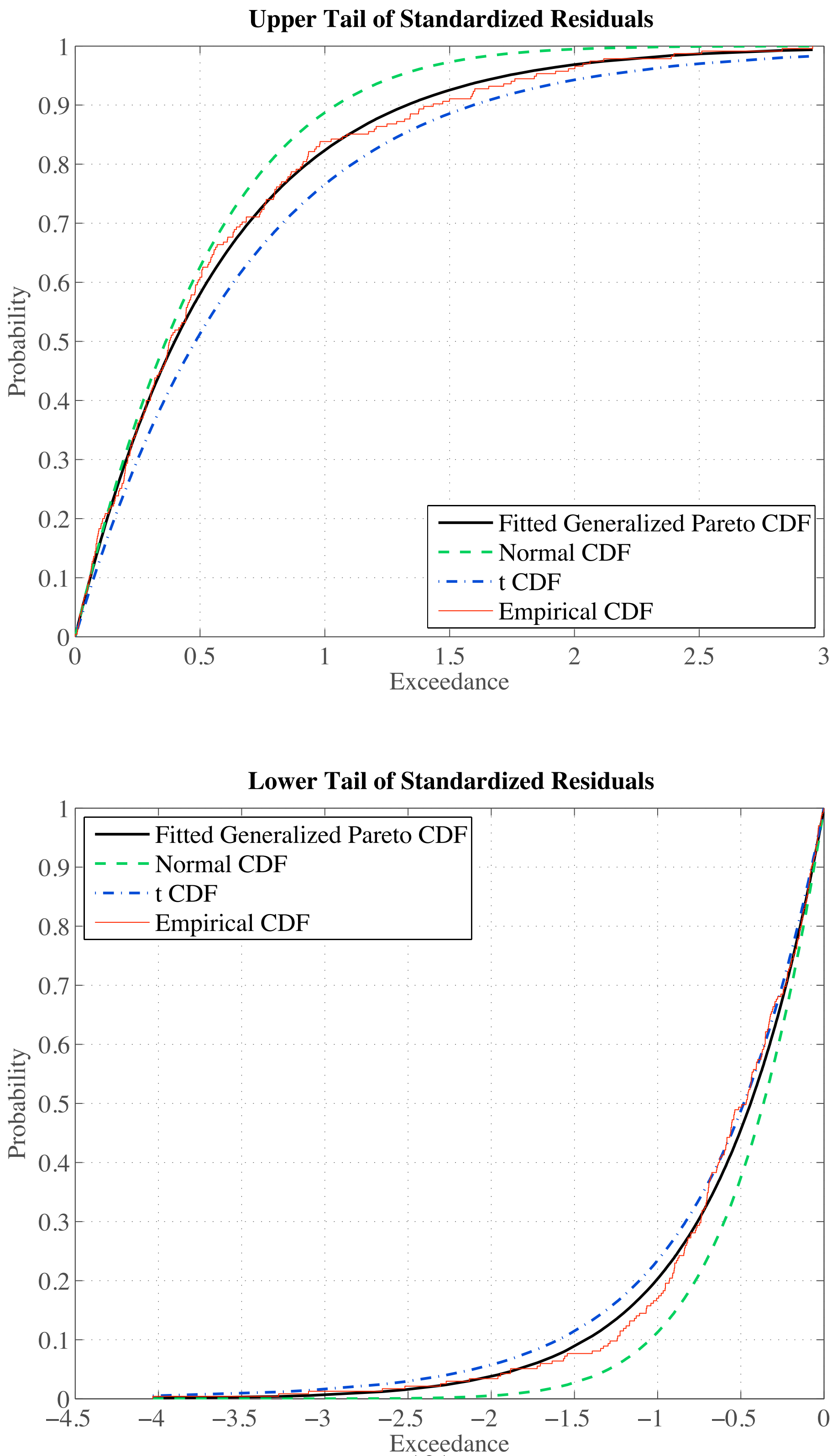

Figure 3.3: **Third principal component.** Upper and lower tails of standardized residuals.

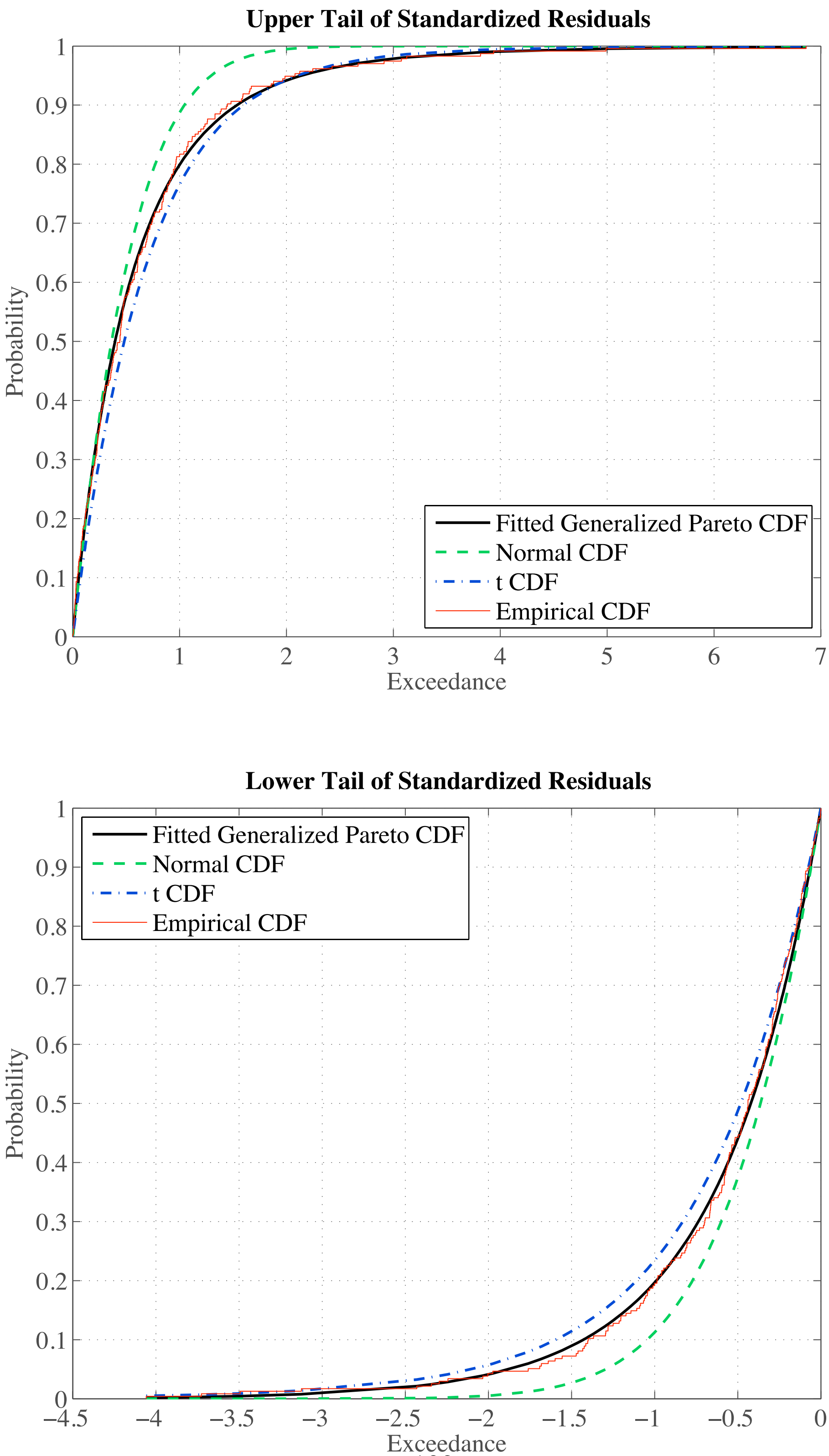

Figure 3.4: **Fourth principal component.** Upper and lower tails of standardized residuals.

Finally, I compute the one-day VaR and ES forecasts for an equally weighted portfolio of currencies using formulas (3.26) and (3.27), for January 1, 2008. The results are summarized in Table 3.7. The forecasts are given for the confidence levels of 90, 95, 99 and 99.9 percent. The values for VaR and ES are reported in percent. The lower- (upper-) tail values are applicable for the losses associated with holding a long (short) position in the portfolio. Evidently, the distributional asymmetry is reflected in pronounced differences between the risk measures in the upper and lower tail.

Table 3.7: **VaR and ES Forecasts**

One-day upper- and lower-tail VaR and ES forecasts for January 1, 2008, for an equally weighted portfolio of currencies and for several confidence levels.

Upper tail				
CL	0.90	0.95	0.99	0.999
VaR	0.5294	0.6947	1.0750	1.6101
ES	0.7284	0.8924	1.2694	1.7994

Lower tail				
CL	0.90	0.95	0.99	0.999
VaR	−0.4212	−0.5923	−0.9203	−1.2495
ES	−0.6790	−0.8251	−1.1026	−1.3743

3.4 Backtesting

Any risk management model needs to be tested before we can successfully apply it in practice. A variety of tests has been proposed to evaluate the accuracy of a VaR model. These tests are constructed to give an assessment of adequacy of the proposed models in predicting the size and frequency of losses. The standard backtests of VaR models compare the VaR forecasts for a given horizon with the actual portfolio losses. In its simplest form, the backtesting procedure consists of calculating the absolute or relative number of times that the actual portfolio returns fall outside the VaR estimate, and comparing that number to the confidence level used.

Model backtesting is also important for financial institutions that are subjected to regulatory requirements. Since the late 1990s, regulatory guidelines require that banks with substantial trading activity have to set aside capital to insure against extreme portfolio losses. The size of the set-aside, or market risk capital requirement, is directly related to a measure of portfolio risk. In most of developed markets, the present regulatory framework follows the recommendations of Basel II, the second of the accords issued by the Basel Committee on Banking Supervision. The purpose of Basel II (initially published in June 2004) and its subsequent amendments was to create an international standard that can be used by national banking regulators. Currently, there are two general methodologies for assessment of market risk capital requirements under Basel II. The first one is the so-called Standardized Approach (SA), and is based on a set of simple rules on how to calculate minimum capital requirements using basic cross-sectional information about the assets in the bank's trading book. The more advanced approach is the Internal Models Method (IMM), which is based on VaR, and – being more precise – typically yields lower capital requirements. Specifically, the minimum capital requirement under IMM is defined as

$$\mathrm{MCR}_t := \max\left\{\mathrm{VaR}_t,\ (M + P_t)\overline{\mathrm{VaR}}\right\} + \mathrm{SRC}_t,$$

where VaR_t is the ten-days-ahead VaR forecast at 99 percent confidence level, $\overline{\mathrm{VaR}}$ is the average of these forecasts over the past 60 trading days, M is a multiplication factor set by national regulators (usually equal to 3), P_t is the penalty associated with the backtesting results, while SRC_t is the specific risk capital charge. The penalty P_t is determined by classifying the number of violations I of one-day 99-percent VaR in the previous 250 trading days into three distinct categories:

- $P_t = 0$, if $I \leq 4$ (green zone);

• $P_t = (I-4)/5$, if $5 \leq I \leq 9$ (yellow zone);
• $P_t = 1$, if $10 \leq I$ (red zone).
Hence, a VaR model with more violations leads to a greater capital requirement.

The Basel II "traffic-light" approach to backtesting represents the only assessment of VaR accuracy prescribed in the current regulatory framework. Although its simple implementation is suitable for informational purposes, this approach merely counts the breaches of the 99-percent confidence level and fails to discard any model that, for example, overestimates the risk, or performs poorly when compared to other confidence levels. The ability of a backtest to discard all the models that systematically overstate as well as understate the risk is known as the unconditional coverage property. Christoffersen (1998) points out that the problem of determining the accuracy of a VaR model can be reduced to the problem of determining whether the sequence of breach counts satisfies both the unconditional coverage and independence. The latter property refers to intuition that the previous history of VaR violations must not convey any information about the future violations.

Some of the earliest VaR backtests proposed in the literature focused on the property of unconditional coverage, that is, whether or not the reported VaR is violated more or less than α percent of the time, where $1-\alpha$ is the confidence level. Kupiec (1995), for example, proposed a proportion of failures (POF) test that examines how many times VaR forecasts are violated over a given span of time. If the number of violations differs significantly from α times the size of the sample, then the accuracy of the underlying risk model is called into question. Using a sample of T observations, Kupiec (1995) test statistic takes the form,

$$\mathrm{POF} := 2\ln\left[\left(\frac{1-\widehat{\alpha}}{1-\alpha}\right)^{T-I(\alpha)}\left(\frac{\widehat{\alpha}}{\alpha}\right)^{I(\alpha)}\right], \tag{3.30}$$

where

$$
\begin{aligned}
\widehat{\alpha} &:= \frac{I(\alpha)}{T}, \\
I(\alpha) &:= \sum_{t=1}^{T} I_t(\alpha),
\end{aligned}
$$

where $I_t(\alpha)$ is an indicator function taking the value one if the actual return at t breaches the forecasted value of VaR for the confidence level determined by α, and zero if it stays within the VaR bounds. Hence, if the proportion of VaR violations, $\widehat{\alpha}$, is exactly equal to α then the POF statistic takes the value zero, indicating no evidence of any inadequacy in the underlying VaR measure. As the proportion of VaR violations differs from α, the POF statistic grows, indicating that the proposed VaR measure either systematically understates or overstates the underlying level of risk. The POF statistic given by (3.30) is a likelihood ratio and hence converges in distribution to a χ^2 with $I(\alpha)$ degrees of freedom.

Figure 3.5 shows the comparison between actual returns (dots) and VaR forecasts (continuous lines) for different confidence levels for an equally weighted portfolio of four currencies, for the period January 1, 2004 – September 30, 2008. The forecasts are computed using the formula (3.26), both for the lower and the upper tail of the return distribution in order to take into account losses both of a long and a short position, respectively. Table 3.8 summarizes the backtesting results, comparing the expected number of violations with the actual ones. The actual violations were compared across different multivariate models (EVT, conditional normality and conditional t-distribution), as well as the univariate historical simulation. The multivariate normal and t models applied here follow the orthogonal GARCH approach of Alexander (2001). In other words, these forecasts were also obtained using the ARMA(1,0)–GJR-GARCH(1,1) filtering of principal components, except that the estimation of the covariance matrices was performed across the entire sample (thereby including both the center and the tails of the distribution) via ML method assuming normally- or t-distributed conditional residuals. The number of violations by quantiles clearly shows that HS markedly deviates from the expected values. The multivariate normal underestimates, while the multivariate t model overestimates the tail risk. At the same time, the multivariate EVT appears to yield much better forecasts. This is also verified formally by means of the Kupiec (1995) test, see Table 3.9. Clearly, all the models give predictions that are within statistically significant bounds for confidence levels of 90 and 95 percent, except for the normal model in the upper tail at 95 percent confidence level. However, the HS model performs poorly at all higher confidence levels, the multivariate t at 99 percent confidence, while the multivariate normal falls short in explaining the upper-tail returns above 99 percent level, and both the upper- and lower-tail extreme returns above 99.9 percent level. On the other hand, VaR forecasts based on the proposed multivariate EVT method violate the corresponding confidence bounds by a number of times that is not statistically different from the expected one. The only exception is perhaps the extreme confidence interval of 99.9 percent, where we observe no violations in the upper tail and one violation in the lower tail, compared to the expectation of 1.239, so for an appropriate sense of statistical significance at these extreme return regions we might need an even longer backtesting sample.

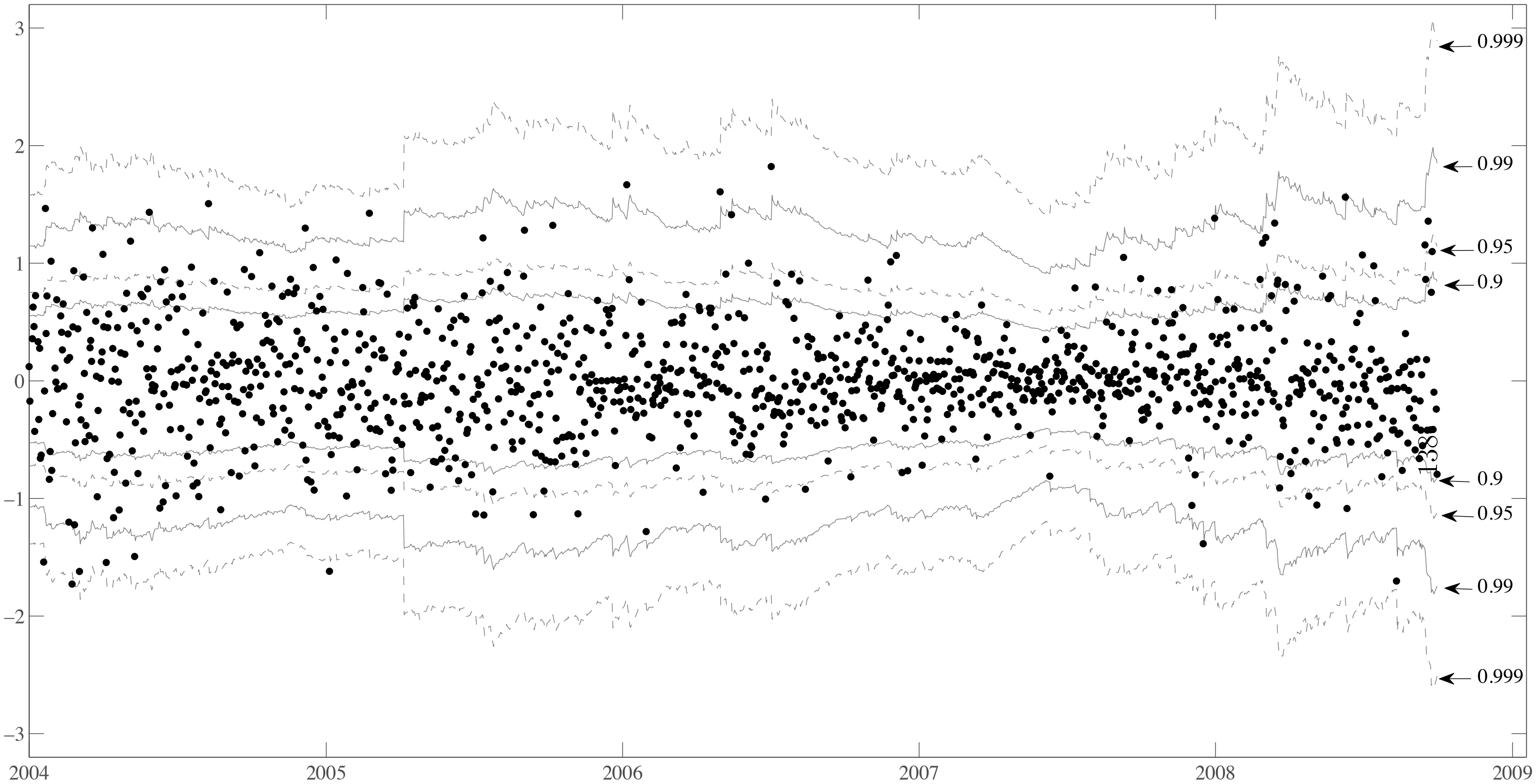

Figure 3.5: **VaR Backtesting of the Multivariate EVT Method.** Actual returns (dots) and VaR forecasts in the upper- and lower-tail at different confidence levels (continuous lines), for an equally weighted portfolio of currencies between January 1, 2004 and September 30, 2008.

Table 3.8: **VaR Backtesting: Violations by Quantiles**

Expected versus the actual number of violations obtained by several models for an equally weighted portfolio of currencies, between January 1, 2004 and September 30, 2008 (a total of 1239 observations).

Upper tail				
Method	Number of violations			
	CL = 0.90	CL = 0.95	CL = 0.99	CL = 0.999
EVT	109	56	12	0
Normal	144	89	31	16
t	121	49	11	0
HS	235	166	70	21
Expected	123.9	61.95	12.39	1.239

Lower tail				
Method	Number of violations			
	CL = 0.90	CL = 0.95	CL = 0.99	CL = 0.999
EVT	88	48	9	1
Normal	118	67	19	8
t	86	40	7	0
HS	219	142	54	28
Expected	123.9	61.95	12.39	1.239

(p-values in parentheses.)

Table 3.9: **VaR Backtesting: Kupiec Test**

The Kupiec (1995) POF test statistics and p-values obtained by several models for an equally weighted portfolio of currencies, between January 1, 2004 and September 30, 2008 (a total of 1239 observations).

Upper tail				
Method	Number of violations			
	CL = 0.90	CL = 0.95	CL = 0.99	CL = 0.999
EVT	2.0665	0.6207	0.0125	2.4792
	($< 10^{-4}$)	($< 10^{-4}$)	($< 10^{-4}$)	$--$
Normal	3.4620	11.0174	19.9238	52.5198
	($< 10^{-4}$)	(0.4564)	(0.0626)	(~ 1.0)
t	0.0759	3.0602	0.1637	2.4792
	($< 10^{-4}$)	($< 10^{-4}$)	(0.2022)	$--$
HS	90.1083	128.6208	129.9539	79.6643
	($< 10^{-4}$)	(0.0142)	(~ 1.0)	(~ 1.0)

Lower tail				
Method	Number of violations			
	CL = 0.90	CL = 0.95	CL = 0.99	CL = 0.999
EVT	12.7273	3.5725	1.0354	0.0494
	($< 10^{-4}$)	($< 10^{-4}$)	(0.0006)	(0.1760)
Normal	0.3167	0.4226	3.0626	16.3572
	($< 10^{-4}$)	($< 10^{-4}$)	($< 10^{-4}$)	(0.9625)
t	14.2719	9.3110	2.8099	2.4792
	($< 10^{-4}$)	($< 10^{-4}$)	(0.0980)	$--$
HS	67.6349	81.0498	77.1940	121.6632
	($< 10^{-4}$)	($< 10^{-4}$)	(0.9791)	(~ 1.0)

(p-values in parentheses.)

By examining a variety of different quantiles instead of a single one, some types of backtests can detect violations of the independence across a range of different VaR levels, while satisfying the unconditional coverage property. A variety of such tests has been proposed during the past decade, and Campbell (2005) gives a good review of these and other backtesting methods. An example of such a test is Pearson's test for goodness of fit. This test is based upon the number of observed violations at a variety of different VaR levels, separated into bins on the unit interval. The Pearson's test statistic is given by

$$Q := \sum_{k=1}^{K} \frac{\left(N_k^{\text{obs}} - N_k^{\text{exp}}\right)^2}{N_k^{\text{exp}}}, \tag{3.31}$$

where N_k^{obs} and N_k^{exp} are, respectively, the observed and the expected number of violations in the k-th bin. The Q statistic converges in distribution to a χ^2 with $K-1$ degrees of freedom, K being the number of bins. The results of the Pearson's test for the currency portfolio are summarized in Table 3.10, for the set of bins given by $\alpha \in [0.00, 0.001) \cup [0.001, 0.01) \cup [0.01, 0.05) \cup [0.05, 0.10) \cup [0.10, 1.00]$. They show that models based on conditionally normal or t-distributed residuals, as well as the HS model, can be rejected in favor of the proposed multivariate EVT alternative.

Table 3.10: **VaR Backtesting: Pearson's Test**

Pearson's test statistics and p-values obtained using VaR forecasts for an equally weighted portfolio of currencies between January 1, 2004 and September 30, 2008 (a total of 1239 observations). The partition of the unit interval used was $\alpha \in [0.00, 0.001) \cup [0.001, 0.01) \cup [0.01, 0.05) \cup [0.05, 0.10) \cup [0.10, 1.00]$.

Method	Lower tail	Upper tail
EVT	0.4170	1.3142
	(0.0189)	(0.1410)
Normal	38.5252	7.5773
	($\sim$ 1.0)	(0.8917)
t	2.2298	2.0934
	(0.3064)	(0.2814)
HS	123.1067	146.7974
	($\sim$ 1.0)	($\sim$ 1.0)

(p-values in parentheses.)

3.5 Conclusion

This paper develops an efficient procedure for estimation of Value at Risk and expected shortfall based on a multivariate extreme value theory approach. The method is based on separate estimations of the univariate EVT model. It works with a set of orthogonal conditional residuals, obtained from the principal components of the joint return series. Autocorrelation, heteroskedasticity and asymmetry that are inherent in the original return series can be removed by assuming an ARMA process for the conditional mean and an asymmetric GARCH process for the conditional variance of the principal components. In this way, we can obtain a set of independent and identically distributed random variables, which is a prerequisite for any univariate EVT approach. The tails of the univariate distributions are modeled by a generalized Pareto distribution of peeks over treshold, while the interiors are fitted with an empirical distribution function. Furthermore, the method can be free of any unnecessary distributional assumption since the estimation of the ARMA-GARCH parameters can be performed via a generalized method of moments. Also, the method is free of estimation of a joint multivariate distribution, which would require a technique such as copula approach with simulations.

As an illustration, the method is applied to a sequence of daily interbank spot exchange rates of Euro, British Pound, Japanese Yen and Swiss Franc with respect to the U.S. Dollar. The forecasts of VaR and ES are backtested through a comparison with the actual losses over an out-of-the-sample period of four years and three quarters. The backtesting results indicate that the proposed multivariate EVT method performs well in forecasting the risk of a portfolio of four currencies. It certainly gives more precise estimate of VaR than the usual methods based on conditional normality, conditional t-distribution or historical simulation, while having the efficiency of an orthogonal GARCH method.

Bibliography

Acerbi, C., and D. Tasche (2002): "Expected Shortfall: A Natural Coherent Alternative to Value at Risk," *Economic Notes*, 31(2), 1–10.

Alexander, C. (2001): "Orthogonal GARCH," in *Mastering Risk*, ed. by C. Alexander, vol. 2, pp. 21–38. Pearson.

Andersen, T. G., L. Benzoni, and J. Lund (2002): "An Empirical Investigation of Continuous-Time Equity Return Models," *Journal of Finance*, 57(3), 1239–1284.

Artzner, P., F. Delbaen, J. M. Eber, and D. Heath (1997): "Thinking Correctly," *Risk*, 10(11), 68–71.

——— (1999): "Coherent Measures of Risk," *Mathematical Finance*, 9(3), 203–228.

Balkema, A. A., and L. de Haan (1974): "Residual Life Time at Great Age," *Annals of Probability*, 2, 792–804.

Ball, C. A., and W. N. Torous (1985): "On Jumps in Common Stock Prices and Their Impact on Call Option Pricing," *Journal of Finance*, 40(1), 155–173.

Bates, D. S. (1988): "The Crash Premium: Option Pricing Under Asymmetric Processes, With Applications to Options on Deutschmark Futures," Rodney L. White Center working paper 36-88, October.

——— (1991): "The Crash of '87: Was It Expected? The Evidence from Options Markets," *Journal of Finance*, 46, 1009–1044.

——— (1996a): "Dollar Jump Fears, 1984–1992: Distributional Abnormalities Implicit in Currency Futures Options," *Journal of International Money and Finance*, 15(1), 65–93.

——— (1996b): "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 S&P 500 Futures Options," *Journal of Econometrics*, 94, 181–238.

Bensalah, Y. (2000): "Steps in Applying Extreme Value Theory to Finance: A Review," Working Paper, Bank of Canada.

Benth, F., L. Ekeland, R. Hauge, and B. Nielsen (2003): "A Note on Arbitrage-Free Pricing of Forward Contracts on Energy Markets," *Applied Mathematical Finance*, 10, 325–336.

Bessembinder, H., J. F. Coughenour, P. J. Seguin, and M. M. Smoller (1995): "Mean Reversion in Equilibrium Asset Prices: Evidence from the Futures Term Structure," *Journal of Finance*, 50(1), 361–375.

Bingham, N., and R. Kiesel (2004): *Risk-Neutral Valuation: Pricing and Hedging of Financial Derivatives*. Springer Finance, London, UK.

Black, F., and M. S. Scholes (1973): "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81(3), 637–654.

Brennan, M. J. (1991): "The Price of Convenience and the Valuation of Commodity Contingent Claims," in *Stochastic Models and Option Values*, ed. by D. Lund, and B. Oksendal. North Holland.

Breuer, T., M. Jandacka, K. Rheinberger, and M. Summer (2008): "Regulatory Capital for Market and Credit Risk Interaction: Is Current Regulation Always Conservative?," Working Paper.

Campbell, S. D. (2005): "A Review of Backtesting and Backtesting Procedures," Finance and Economics Discussion Series, Federal Reserve Board, Washington, D.C.

Cao, M. (2001): "Systematic Jump Risks in a Small Open Economy: Simultaneous Equilibrium Valuation of Options on the Market Portfolio and the Exchange Rate," *Journal of International Money and Finance*, 20(2), 191–218.

Casassus, J., and P. Collin-Dufresne (2005): "Stochastic Convenience Yield Implied from Commodity Futures and Interest Rates," *Journal of Finance*, 60(5).

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, Pennsylvania State University.

Chernov, M., and E. Ghysels (2000): "A Study Towards a Unified Approach to the Joint Estimation of Objective and Risk Neutral Measures for the Purpose of Options Valuation," *Journal of Financial Economics*, 56, 407–458.

Christoffersen, P. (1998): "Evaluating Interval Forecasts," *International Economic Review*, 39, 841–862.

Clewlow, L., and C. Strickland (2000): "Energy Derivatives: Pricing and Risk Management," Working Paper, Lacima Group.

Clewlow, L., C. Strickland, and V. Kaminski (2001): "Extending Mean-Reversion Jump Diffusion," Energy Power Risk Management, Risk Waters Group (February).

Cortazar, G., and E. S. Schwartz (1994): "The Evaluation of Commodity Contingent Claims," *Journal of Derivatives*, 1, 27–39.

——— (2003): "Implementing a Stochastic Model for Oil Futures Prices," *Energy Economics*, 25, 215–238.

Cox, J. C., J. E. Ingersoll, and S. A. Ross (1981): "The Relation Between Forward Prices and Futures Prices," *Journal of Financial Economics*, 9, 321–346.

——— (1985): "A Theory of the Term Structure of Interest Rates," *Econometrica*, 53, 385–407.

Craine, R., L. A. Lochstoer, and K. Syrtveit (2000): "Estimation of a Stochastic-Volatility Jump-Diffusion Model," Working Paper, University of California at Berkeley.

Crosby, J. (2005): "A Multi-Factor Jump-Diffusion Model for Commodities," Working Paper.

da Silva, L. C. A., and B. V. de Melo Mendez (2003): "Value-at-Risk and Extreme Returns in Asian Stock Markets," *International Journal of Business*, 8(1), 17–40.

Danielsson, J., and C. de Vries (1997): "Value at Risk and Extreme Returns," FMG Discussion paper No. 273, Financial Markets Group, London School of Economics.

Davison, A. C., and R. L. Smith (1990): "Models for Exceedances over High Tresholds," *Journal of the Royal Statistical Society B*, 52, 393–442.

Deng, S. (1998): "Stochastic Models of Energy Commodity Prices and Their Applications; Mean Reversion with Jumps and Spikes," Univ. of California at Berkeley Working Paper No. 94720.

DIAS, M. A. G., AND K. M. C. ROCHA (2001): "Petroleum Concessions with Extendible Options using Mean Reversion with Jumps to Model Oil Prices," Working Paper.

DOFFOU, A., AND J. E. HILLIARD (2001): "Pricing Currency Options Under Stochastic Interest Rates and Jump-Diffusion Processes," *Journal of Financial Research*, 24(4), 565–586.

DOWD, K. (1998): *Beyond Value at Risk: The New Science of Risk Management.* John Wiley & Sons, Chichester.

DUAN, J.-C. (1994): "Maximum Likelihood Estimation Using Price Data of the Derivative Contract," *Mathematical Finance*, 4(2), 155–167.

DUFFIE, D. (1989): *Futures Markets.* Prentice Hall, Englewood Cliffs, NJ.

——— (1992): *Dynamic Asset Pricing Theory.* Princeton University Press, Princeton, NJ.

DUFFIE, D., J. PAN, AND K. SINGLETON (2000): "Transform Analysis and Asset Pricing for Affine Jump-Diffusions," *Econometrica*, 68(6), 1343–1376.

DUFFIE, D., AND K. J. SINGLETON (1993): "Simulated Moments Estimation of Markov Models of Asset Prices," *Econometrica*, 61, 929–952.

ELLIOT, R. J., AND P. E. KOPP (2005): *Mathematics of Financial Markets.* Springer Finance, London, UK.

ELLSBERG, D. (1961): "Risk, Ambiguity, and the Savage Axioms," *Quarterly Journal of Economics*, 75, 643–669.

EMBRECHTS, P., C. KLÜPPELBERG, AND T. MIKOSCH (2008): *Modeling Extremal Events for Insurance and Finance.* Springer, New York.

FISHER, R. A., AND L. H. C. TIPPETT (1928): "Limiting Forms of the Frequency Distribution of the Largest or Smallest Member of a Sample," *Proceedings of the Cambridge Philosophical Society*, 24, 180–190.

Gallant, A. R., and J. R. Long (1997): "Estimating Stochastic Differential Equation Efficiently by Minimum Chi-Square," *Biometrika*, 84, 125–142.

Gallant, A. R., and D. W. Nychka (1987): "Semi-Nonparametric Maximum Likelihood Estimation," *Econometrica*, 55, 363–390.

Gallant, A. R., and G. Tauchen (1996): "Which Moments to Match?," *Econometric Theory*, 12, 657–681.

Gibson, R., and E. S. Schwartz (1990): "Stochastic Convenience Yield and the Pricing of Oil Contingent Claims," *Journal of Finance*, 45(3), 959–976.

Glosten, L. R., R. Jagannathan, and D. E. Runkle (1993): "On the Relation between Expected Value and the Volatility of the Nominal Excess Return on Stocks," *Journal of Finance*, 48, 1779–1801.

Gnedenko, B. V. (1943): "Sur la Distribution limite du Terme maximum d'une Série aléatoir," *Annals of Mathematics*, 44, 423–453.

Harmantzis, F., Y. Chien, and L. Miao (2005): "Empirical Study of Value-at-Risk and Expected Shortfall as a Coherent Measure of Risk," *Journal of Risk Finance*, 7(2), 117–135.

Harrison, J. M., and S. R. Pliska (1981): "Martingales and Stochastic Integrals in the Theory of Continuous Trading," *Stochastic Processes and their Applications*, 11, 215–260.

Harvey, A. C. (1994): *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press.

Heath, D., R. Jarrow, and A. Morton (1992): "Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation," *Econometrica*, 60(4), 77–105.

Heston, S. L. (1993): "A Closed-form Solution for Options with Stochastic Volatility and Applications to Bond and Currency Options," *Review of Financial Studies*, 6(2), 327343.

Hilliard, J. E., and J. Reis (1998): "Valuation of Commodity Futures and Options Under Stochastic Convenience Yield, Interest rates, and Jump Diffusions in the Spot," *Journal of Financial and Quantitative Analysis*, 33(1), 61–86.

Jiang, G. J. (1998): "Jump Diffusion Model of Exchange Rate Dynamics – Estimation via Indirect Inference," Working Paper.

Jiang, G. J., and R. C. A. Oomen (2008): "Testing for Jumps when Asset Prices are Observed with Noise–A "Swap Variance" Approach," *Journal of Econometrics*, 144(2), 352–370.

Johannes, M. (2004): "The Statistical and Economic Role of Jumps in Continuous-Time Interest Rate Models," *Journal of Finance*, 59(1), 227–260.

Jondeau, E., and M. Rockinger (2003): "Testing for Differences in the Tails of Stock Market Returns," *Journal of Empirical Finance*, 10, 559–581.

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

——— (1996): *Value at Risk: The New Benchmark for Controlling Market Risk*. Irwin, Chicago.

Knight, F. (1921): *Risk, Uncertainty and Profit*. Houghton, Mifflin, Boston.

Kupiec, P. (1995): "Techniques for Verifying the Accuracy of Risk Management Models," *Journal of Derivatives*, 3, 73–84.

Lei, M., and G. Fox (2004): "Operating Options and Commodity Price Processes," Working Paper, University of Guelph.

Liu, J., J. Pan, and T. Wang (2005): "An Equilibrium Model of Rare-Event Premia and Its Implications for Option Smirks," *Review of Financial Studies*, 18(1), 131–164.

Lo, A. W. (1988): "Maximum Likelihood Estimation of Generalized Itô Processes with Discretely Sampled Data," *Econometric Theory*, 4, 231–247.

Maheu, J., and T. McCurdy (2006): "Modeling Foreign Exchange Rates with Jumps," Working Paper, University of Toronto.

Marinelli, C., S. d'Addona, and S. T. Rachev (2007): "A Comparison of Some Univariate Models for Value-at-Risk and Expected Shortfall," *International Journal of Theoretical and Applied Finance*, 10(6), 1–33.

McNeil, A., and R. Frey (2000): "Estimation of Tail Related Risk Measure for Heteroscedastic Financial Time Series: An Extreme Value Approach," *Journal of Empirical Finance*, 7, 271–300.

McNeil, A. J. (1997): "Estimating the Tails of Loss Severity Distribution Using Extreme Value Theory," *Theory ASTIN Bulletin*, 27(1), 1117–1137.

——— (1999): "Extreme Value Theory for Risk Managers," Working Paper, ETH Zürich.

Merton, R. C. (1976): "Option Pricing when Underlying Stock Returns are Discontinuous," *Journal of Financial Economics*, 3, 125–144.

——— (1990): *Continuous-Time Finance*. Blackwell, Oxford.

Miltersen, K. R. (2003): "Commodity Price Modelling that Matches Current Observables: A New Approach," *Quantitative Finance*, 3, 51–58.

Miltersen, K. R., and E. S. Scwartz (1998): "Pricing of Options on Commodity Futures with Stochastic Term Structure of Convenience Yield and Interest Rates," *Journal of Financial and Quanititative Analysis*, 33(1), 33–59.

Nelson, D. B. (1991): “Conditional Heteroskedasticity in Asset Returns: A New Approach,” *Econometrica*, 59, 347–370.

Newey, W. K., and D. McFadden (1994): “Large Sample Estimation and Hypothesis Testing,” in *Handbook of Econometrics*, ed. by R. F. Engle, and D. McFadden, vol. 4, pp. 2111–2245, Amsterdam. North Holland.

Nyström, K., and J. Skoglund (2002a): “A Framework for Scenario-Based Risk Management,” Working Paper, Swedbank.

——— (2002b): “Univariate Extreme Value Theory, GARCH and Measures of Risk,” Working Paper, Swedbank.

Olsen, R. B., U. A. Müller, M. M. Dacorogna, O. V. Pictet, R. R. Davé, and D. M. Guillaume (1997): “From the Bird’s Eye to the Microscope: A Survey of New Stylized Facts of the Intra-daily Foreign Exchange Markets,” *Finance and Stochastics*, 1(2), 95–129.

Pan, J. (2002): “The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time-Series Study,” *Journal of Financial Economics*, 63, 3–50.

Picklands, J. (1975): “Statistical Inference Using Extreme Order Statistics,” *Annals of Statistics*, 3, 119–131.

Pilipovic, D. (1998): *Energy Risk*. McGraw-Hill, New York.

Pindyk, R. S. (2001): “The Dynamics of Commodity Spot and Futures Markets: A Primer,” Working Paper, MIT.

Pindyk, R. S., and D. L. Rubinfeld (1991): *Econometric Models and Economic Forecasts*. McGraw-Hill, New York.

Ross, S. A. (1995): “Hedging Long Run Commitments: Exercises in Incomplete Market Pricing,” Working Paper.

Rougier, G. (1992): "An Optimal Price Index for Stock Index Futures Contracts," *Journal of Futures Markets*, 16, 189–199.

Schwartz, E. S. (1997): "The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging," *Journal of Finance*, 52(3), 923–973.

Skoglund, J. (2001): "A Simple Efficient GMM Estimator of GARCH Models," Working paper series in economics and finance, Stockholm School of Economics.

Smith, R. L. (2000): "Measuring Risk with Extreme Value Theory," in *Risk Management: Theory and Practice*. Cambridge University Press.

Taylor, S. J. (1986): *Modeling Financial Time Series*. John Wiley.

Vasicek, O. (1977): "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics*, 5, 177–188.

von Mises, R. (1936): "La Distribution de la plus grande de n Valeurs," in *Selected Papers II, 1954*, pp. 271–294, Providence, RI. American Mathematical Society.