Université de Montréal

Trois essais sur la liquidité : ses effets sur les primes de risque, les anticipations et l'asymétrie des risques financiers.

par

Jean-Sébastien Fontaine

Département des Sciences Économiques
Faculté des Sciences Humaines

Thèse présentée à la Faculté en vue de l'obtention du grade de Doctorat en Sciences Économiques

Mai 2009

Université de Montréal
Faculté des études supérieures

Cette thèse intitulée

Trois essais sur la liquidité : ses effets sur les primes de risque, les anticipations et l'asymétrie des risques financiers.

présentée par

Jean-Sébastien Fontaine

a été évaluée par un jury composé de

Silvia Gonçalvez
président-rapporteur

René Garcia
directeur de recherche

Jean Boivin
membre du jury

Monika Piazzesi
examinatrice externe

William McCausland
Représentant de l'examinatrice externe

Thomas Legrand
Représentant du doyen

Résumé

Le premier des trois essais qui composent cette thèse démontre empiriquement que la contrainte de crédit à laquelle font face les intermédiaires financiers affecte leur capacité à supporter un marché liquide. Dans le cadre des marchés américains, cet essai mesure les variations du prix de cette contrainte et les relient aux variations des primes de risque dans différents marchés. L'évidence démontre que cette contrainte financière induit un facteur de risque agrégé et cause des variations importantes et communes des primes de liquidité à travers les marchés monétaires et obligataires.

Le second essai s'intéresse au contenu informatif des taux d'intérêts et des contrats à terme liés aux décisions futures de la Réserve Fédérale américaine. Je développe un modèle joint de la politique monétaire, de la structure des taux d'intérêts ainsi que des contrats à termes. En particulier, le modèle permet une prime de liquidité différente dans chaque marché. Les résultats montrent que cette approche permet d'identifier plus précisément les anticipations incorporées dans le prix de ces actifs et, ainsi, d'améliorer la prévision des décisions monétaires. L'identification, par le modèle, des variations dans la prime de liquidité contribuent pour beaucoup à ces résultats.

Finalement, le dernier essai s'intéresse à l'importance de l'asymétrie des rendements boursiers. En particulier pour expliquer les écarts systématiques observés entre la volatilité des rendements futurs et celle implicite dans les prix d'options liées aux rendements futurs. Le modèle prévoit, et les résultats supportent cette conclusion, que tenir compte de l'asymétrie anticipée, telle que mesurée à partir de prix d'options, apporte une contribution à la prévision de la prime de risque et à notre capacité à mesurer et couvrir les risques reliés aux options. Les liens entre l'asymétrie des rendements et l'écart de volatilité sont des élément-clés pour notre compréhension des rendements boursiers.

Mots-clés : Économie financière, Économétrie financière, prime de liquidité, taux d'intérêts, prix d'option, asymétrie et prime de risque.

Abstract

The first essay of this thesis provides evidence that the credit constraint of financial intermediaries affect their ability to support liquid markets. The essay studies the case of US markets and provides a direct measurement, based on observed asset prices, of that constraint's price variations. The evidence shows that these variations represent an aggregate risk factor and induce large common variations of the liquidity premia across money markets and bond markets.

The second essay considers the information content of interest rates and futures contracts linked to future policy decision by the US Federal Reserve. I develop a joint model of the monetary policy response function, of interest rates and of futures rates. In particular, the model allows for different liquidity premia across the two markets. Empirically, this approach identifies more precisely the expectations of future policy implicit in observed prices and, hence, enhances our ability to forecast monetary decisions. Measuring the liquidity premium contributes significantly to these improvements.

Finally, the last essay studies the importance of the skewness of stock index returns. In particular, we analyze the role of skewness in the systematic spread between the observed volatility of equity returns and the volatility implicit in index option prices. The model predicts and the evidence supports that conditioning on implied skewness, as measured from option prices, improves our ability to predict the risk premium and our ability to measure and hedge option-related risk exposures. We argue that the linkages between skewness and the volatility spread are key to our understanding of the equity premium.

Keywords : Financial Economics, Financial Econometrics, Liquidity Premium, interest rates, option prices, skewness and risk premium.

Table des matières

Table des figures

Liste des tableaux

Liste des abbréviations

bps	basis points
BSM	Black-Scholes-Merton
EKF	Extended Kalman Filter
Fed	Federal Reserve
FOMC	Federal Open Market Committee
HG	Homoscedastic Gamma
LIBOR	London Inter-Bank Offered Rate
P-BSM	Practitioner's Black-Scholes-Merton
P-HG	Practitioner's Homoscedastic Gamma
SDF	Stochastic Discount Factor
SDG	Standardized Gamma
UKF	Unscented Kalman Filter

À Lise et Jean-Yves, pour leur support et m'avoir transmis le désir de savoir.

À Julie, qui partage ma vie, et les aléas de la recherche.

À Maya.

Je remercie J.C.R. Rowley de l'université Mcgill qui m'a appris que la recherche est d'abord une affaire de passion plutôt qu'une affaire de tête. Je remercie aussi mes professeurs à l'université de Montréal. En particulier Nour Meddahi. Cette thèse n'aurait pas été possible sans son intervention et ses exhortations. Finalement, je ne peux exprimer suffisamment ma reconnaissance envers René Garcia pour la confiance qu'il m'a accordée et son support constant. Il a été un mentor, et il est pour moi un exemple.

Le Centre interuniversitaire de recherche en analyse des organisations [CIRANO] et le Centre interuniversitaire de recherche en économie quantitative [CIREQ] fournissent un environnement de recherche unique grâce à leur soutien financier et matériel ainsi qu'aux nombreuses opportunités offertes au travers des conférences internationales qu'ils supportent à Montréal. Finalement, je remercie l'Institut de Mathématique de Montréal et la Banque Laurentienne pour leur soutien financier et leur implication pour la recherche.

Apport des coauteurs

Chapitre 1 : Bond Liquidity Premia

Cet article a été écrit en collaboration avec mon superviseur René Garcia. Bien que j'ai été maître d'oeuvre pour cet article, ce dernier a joué un rôle crucial lors de certaines étapes clés. Nos contributions à l'élaboration du projet de recherche, la spécification de la méthodologie employée, et la rédaction de l'article sont équivalentes.

Chapitre 2 : The Term Structure Of Monetary Expectations

Cet article a été rédigé sans coauteur.

Chapitre 3 : The Equity Premium Premium And The Volatility Spread

Cet article a été écrit en collaboration avec Bruno Feunou et Roméo Tédongap. Bien que j'ai été maître d'oeuvre pour la rédaction de cet article, ces derniers ont joué un rôle crucial lors de certaines étapes clés. Nos contributions à l'élaboration du projet de recherche, la spécification de la méthodologie employée et l'interprétation des résultats sont équivalentes.

Chapitre 1

Bond Liquidity Premia

Bond Liquidity Premia

Jean-Sébastien Fontaine
Université de Montréal and CIREQ
René Garcia
EDHEC Business School

Abstract

Recent asset pricing models of limits to arbitrage emphasize the role of funding conditions faced by financial intermediaries. In the US, the repo market is the key funding market. Then, the premium of on-the-run U.S. Treasury bonds should share a common component with risk premia in other markets. This observation leads to the following identification strategy. We measure the value of funding liquidity from the cross-section of on-the-run premia by adding a liquidity factor to an arbitrage-free term structure model. As predicted, we find that funding liquidity explains the cross-section of risk premia. An increase in the value of liquidity predicts lower risk premia for on-the-run *and* off-the-run bonds but higher risk premia on LIBOR loans, swap contracts and corporate bonds. Moreover, the impact is large and pervasive through crisis and normal times. We check the interpretation of the liquidity factor. It varies with transaction costs, S&P500 valuation ratios and aggregate uncertainty. More importantly, the liquidity factor varies with narrow measures of monetary aggregates and measures of bank reserves. Overall, the results suggest that different securities serve, in part, and to varying degrees, to fulfill investors' uncertain future needs for cash depending on the ability of intermediaries to provide immediacy.

JEL Classification: E43, H12.

We thank Greg Bauer, Antonio Diez, Darrell Duffie, Thierry Foucault, Francis Longstaff, Albert Menkveld, Monika Piazzesi, Robert Rasche, Jose Sheinkman, participants at the Econometric Society Summer Meeting (2007) and European Meeting (2007), Canadian Economic Association (2007), International Symposium on Financial Engineering and Risk Management (2007), the Risk Management Institute (2009), University of Lugano (2009), University of Piraeus (2009), University College Dublin (2009), WFA (2009) and EFA (2009). The first author gratefully acknowledges support from the IFM^2 and the Banque Laurentienne. The second author is a research fellow at CIRANO and CIREQ. He gratefully acknowledges support from FQRSC, SSHRC, MITACS, Hydro-Québec, and the Bank of Canada. *Correspondence* [redacted]

> "... a part of the interest paid, at least on long-term securities, is to be attributed to uncertainty of the future course of interest rates."
> (p.163)
>
> "... the imperfect 'moneyness' of those bills which are not money [...] causes the trouble of investing in them and [causes them] to stand at a discount."
> (p.166)
>
> "... In practice, there is no rate so short that it may not be affected by speculative elements; there is no rate so long that it may not be affected by the alternative use of funds in holding cash."
> (p.166)

John R. Hicks, Value and Capital, 2nd edition, 1948.

Introduction

Bond traders know very well that liquidity affects asset prices. One prominent case is the on-the-run premium, whereby the most recently issued (on-the-run) bonds sell at a premium relative to seasoned (off-the-run) bonds with similar coupons and maturities. Moreover, systematic variations in liquidity sometimes drive interest rates across several markets. A case in point occurred around the Federal Open Market Committee [FOMC] decision, on October 15, 1998, to lower the Federal Reserve funds rate by 25 basis points. In the meeting's opening, Vice-Chairman McDonough, of the New York district bank, noted increases in the spread between the on-the-run and the most recent off-the-run 30-year Treasury bonds (0.05% to 0.27%), the spreads between the rate on the fixed leg of swaps and Treasury notes with two years and ten years to maturity (0.35% to 0.70%, and 0.50% to 0.95%, respectively), the spreads between Treasuries and investment-grade corporate securities (0.75% to 1.24%), and finally between Treasuries and mortgage-backed securities (1.10% to 1.70%). He concluded that we were seeing a run to quality and a serious drying up of liquidity[1]. These events attest to the sometimes dramatic impact of liquidity seizures[2].

[1] Minutes of the Federal Open Market Committee, October 15, 1998 conference call. See http://www.federalreserve.gov/Fomc/transcripts/1998/981015confcall.pdf.

[2] The liquidity crisis of 2007-2008 provides another example. Facing sharp increases of interest rate spreads in most markets, the Board approved reduction in discount rate,

A common explanation for that and the more recent market turmoil is based on a common wealth shock to capital-constrained intermediaries or speculators (Shleifer and Vishny (1997), Kyle and Xiong (2001), Gromb and Vayanos (2002)). Intuitively, lower wealth hinders the ability to pursue quasi-arbitrage opportunities across markets. In practice, the repo market is the key market where investment banks, hedge funds and other speculators obtain the marginal funds for their activities and manage their leveraged exposure to risk (Adrian and Shin (2008)). Then, the risk premia for each market intermediated by a common set of intermediaries share a component measuring tightness in the funding market (Brunnermeier and Pedersen (2008), Krishnamurthy and He (2008)). This paper tests the implication that tightness of funding conditions in repo markets should be reflected in risk premia across financial markets.

We introduce liquidity as an additional factor in an otherwise standard term structure model. Indeed, the modern term structure literature has not recognized the importance of aggregate liquidity for government yields. We extend the no-arbitrage dynamic term structure model of Christensen et al. (2007) [CDR, hereafter] allowing for liquidity[3] and we extract a common factor driving on-the-run premia across maturities. Identification of the liquidity factor is obtained by estimating the model from a panel of pairs of U.S. Treasury securities where each pair has similar cash flows but different ages. This sidesteps credit risk issues and delivers direct estimates of funding liquidity value: it isolates price differences that can be attributed to liquidity. A recent empirical literature suggests that liquidity is priced on bond markets[4] but these empirical investigations are limited to a single market. Moreover, none consider the role of funding constraints.

Our main contribution is precisely to show that funding liquidity is an aggregate risk factor that drives a substantial share of risk premia across interest rate markets. In particular, we document large variations in the liquidity premium of U.S. Treasury bonds. By construction, an increase in the liquidity factor is associated with lower expected returns for on-the-run

target Federal Funds rate as well as novel policy instruments to deal with the ongoing liquidity crisis.

[3]This model captures parsimoniously the usual level, slope and curvature factors, while delivering good in-sample fit and forecasting power. Moreover, the smooth shape of Nelson-Siegel curves identifies small deviations, relative to an idealized curve, which may be caused by variations in market liquidity.

[4]See Longstaff (2000) for evidence that liquidity is priced for short-term U.S. Treasury security and Longstaff (2004) for U.S. Treasury bonds of longer maturities. See Collin-Dufresne et al. (2001), Longstaff et al. (2005), Ericsson and Renault (2006), Nashikkar and Subrahmanyam (2006) for corporate bonds.

bonds. What we show is that the risk premium of any U.S. Treasury bonds also decreases substantially. On the other hand, tight funding conditions raise the risk premium implicit in LIBOR rates, swap rates and corporate bond yields. This pattern is consistent with accounts of flight-to-quality but the relationship is pervasive even in normal times. This adds considerably to the existing evidence pointing toward the importance of funding liquidity as an aggregate risk factor. Moreover, it suggests that different securities serve, in part and to varying degrees, to fulfill investors uncertain future needs for cash.

We estimate the model and obtain a measure of funding liquidity value from a sample of end-of-month bond prices running from December 1985 until the end of 2007. Hence, our results cannot be attributed to the extreme influence of 2008. In a concluding section, we repeat the estimation including 2008 and find, not unexpectedly, that importance of funding liquidity increases. Our empirical findings can be summarized as follows. Panel (a) of Figure 1.1 presents the measure of funding liquidity value. Clearly, it exhibits significant variations through normal and crisis periods. In particular, the stock crash of 1987, the Mexican Peso devaluation of December 1994, the LTCM failure of 1998 and the recent liquidity crisis are associated with peaks in investors' valuation of the funding liquidity of on-the-run bonds. The relationship with the risk premium of government bonds is illustrated in Figure 1.2. Panel (a) compares the funding liquidity factor with annual excess returns on a 2-year to maturity off-the-run bond. Clearly, an increase in the value of liquidity predicts lower expected excess returns and, thus, higher current bond prices. For that maturity, a one-standard deviation shock to liquidity predicts a decrease in excess returns of 85 basis points [bps] compared to an average excess returns of 69 bps. We obtain similar results using different maturities or investment horizons. Intuitively, while an off-the-run bond may be less liquid relative to an on-the-run bond with similar characteristics, it is still viewed as a liquid substitute. In particular, it can still be quickly converted into cash, at low costs, via the funding market. .

Next, we consider the predictive power of funding liquidity for the risk premium on short-term Eurodollar loans. Panel (b) of Figure 1.2 shows that variations of LIBOR excess returns are positively linked to variations of funding liquidity. The relationship is significant, both statistically and economically. Consider excess returns from borrowing at the risk-free rate for 12 months and rolling a 3-month LIBOR loans. On average, returns from this strategy are not statistically different than zero since the higher term premium on the borrowing leg compensates for the 3-month LIBOR spread earned on the lending leg. However, following a one-standard deviation shock to the funding liquidity factor, rolling excess returns increase by 42 bps. We reach similar conclusions using LIBOR spreads as ex-ante measures of risk premium. The effect of funding liquidity also extends to swap markets. Panel (d) compares the liquidity factor with the spread, above the par Treasury yield, of a swap contract with 5 years to maturity. We find that a shock to funding liquidity predicts an increase of 6 bps the 5-year swap spread. This is economically significant given the higher sensitivity (i.e. duration) of this contract value to changes in yields. In each regression, we control for variations in the level and shape of the term structure of Treasury yields. The marginal contribution of liquidity to the predictive power is high.

Finally, we consider a sample of corporate bond spreads from the NAIC. We find that the impact of liquidity is significant and follows a flight-to-quality pattern across ratings. For bonds of the highest credit quality, spreads decrease, on average, following a shock to the funding liquidity factor. In contrast, spreads of bonds with lower ratings increase. We also compute excess returns on AAA, AA, A, BBB and High Yield Merrill Lynch corporate bond indices (see Figure 1.3) and reach similar conclusions. Bonds with high credit ratings were perceived to be liquid substitutes to government securities and offered lower risk premium following increases of the liquidity factor. This corresponds to an average effect through our sample, the recent events suggests that this is not always the case.

These results raise the all important issue of identifying macroeconomic drivers of the liquidity factor. Can we characterize the aggregate liquidity premium in terms of economic state variables? First, consistent with theory, our liquidity factor varies with measures of transaction costs on the bond market. Second, we find that funding liquidity is linked to stock market valuation ratios and option-implied volatility from S&P 500 index options. These results support empirically the link between conditions on the funding market, the ability of intermediaries to provide liquidity and the level and risk of aggregate wealth. Most importantly, we find that measures of changes in monetary aggregates and changes in bank reserves are key determinants of our liquidity measure. These findings support our interpretation of the liquidity factor as a measure of conditions on the funding market. This provides a third important empirical contribution.

Related Literature

A few empirical papers document the effects of intermediation constraints in specific markets[5] but we differ in significant ways from existing work. First, we measure the effect of intermediation constraints directly from observed prices rather than quantities. Prices aggregate information about and anticipations of intermediaries wealth, their portfolios and the margins they face. Second, we study a cross-section of money-market and fixed-income securities, providing evidence that funding constraints should be thought as an aggregate risk factor driving liquidity premia across markets.

We introduce a measure of funding liquidity *value* based on the higher valuation of on-the-run bonds relative to off-the-run bonds .[6] The on-the-run liquidity premium was first documented by Warga (1992). Amihud and Mendelson (1991) and, more recently, Goldreich et al. (2005) confirm the link between the premium and expected transaction costs. Duffie (1996) provides a theoretical channel between on-the-run premia and lower financing costs on the repo market. Vayanos and Weill (2006) extend this view and model search frictions in both the repo and the cash markets explicitly.[7] The key frictions differentiating bonds with identical cash flows lies in their segmented funding markets. The link between the repo market and the on-the-run premium has been confirmed empirically. (See Jordan and Jordan (1997), Krishnamurthy (2002), Buraschi and Menini (2002) and Cheria et al. (2004).)

We differ from the modern term structure literature in two significant ways. First, the latter focuses almost exclusively on bootstrapped zero-coupon yields[8]. This approach is convenient because a large family of models delivers zero-coupon yields which are linear in the state variables (see Dai and Singleton (2000)). However, we argue that pre-processing the data wipes out the most accessible evidence on liquidity, that is the on-the-run premium.

[5]See Froot and O'Connell (2008) for catastrophe insurance, Gabaix et al. (2009) for mortgage-backed securities, Gârleanu et al. (2009) for index options and Adrian et al. (2009) for exchange rates.

[6]The U.S Treasury recognizes and takes advantages of this price differential: "In addition, although it is not a primary reason for conducting buy-backs, we may be able to reduce the government's interest expense by purchasing older, "off-the-run" debt and replacing it with lower-yield "on-the-run" debt." [Treasury Assistant Secretary for financial markets Lewis A. Sachs, Testimony before the House Committee on Ways and Means].

[7]Kiyotaki and Wright (1989) introduced search frictions in monetary theory and Shi (2005) extends this framework to include bonds. See Shi (2006) for a review. Search frictions can also rationalize the spreads between bid and ask prices offered by market intermediaries (Duffie et al. (2005)).

[8]The CRSP data set of zero-coupon yields is the most commonly used. It is based on the bootstrap method of Fama and Bliss (1987) [FB].

Therefore, we use coupon bond prices directly. However, the state space is no longer linear and we handle non-linearities with the Unscented Kalman Filter [UKF], an extension of the Kalman Filter for non-linear state-space systems (Julier et al. (1995) and Julier and Uhlmann (1996)). We first estimate a model without liquidity and, notwithstanding differences in data and filtering methodologies, our results are consistent with CDR. However, pricing errors in this standard term structure model reveals systematic differences within pairs, correlated with ages. Estimation of the model with liquidity produces a persistent factor capturing differences between prices of recently issued bonds and prices of older bonds. The on-the-run premium increases with maturity but decays with the age of a bond. These new features complete our contributions to the modeling of the term structure of interest rates in presence of a liquidity factor.

We also differ from the recent literature using a reduced-form approach that model a convenience yield in interest rate markets (Duffie and Singleton (1997)). A one-factor model of the convenience yield cannot match the pattern of on-the-run premia across maturities. Moreover, the link between the premium and the age of a bond cannot be captured in a frictionless arbitrage-free model. Still, Grinblatt (2001) argues that the convenience yields of U.S. Treasury bills can explain the U.S. Dollar swap spread. Recently, Liu et al. (2006) and Fedlhütter and Lando (2007) evaluate the relative importance of credit and liquidity risks in swap spreads. Other empirical investigations are related to our work. Jump risk (Tauchen and Zhou (2006)) or the debt-gdp ratio (Krishnamurthy and Vissing-Jorgensen (2007)) have been proposed to explain the non-default component of corporate spreads. Finally, Pastor and Stambaugh (2003) and Amihud (2002) provide evidence of a liquidity risk factor in expected stock returns.

The link between interest rates and aggregate liquidity is supported elsewhere in the theoretical literature. Svensson (1985) uses a cash-in-advance constraint in a monetary economy. Bansal and Coleman (1996) allow government bonds to back checkable accounts and reduced transaction costs in a monetary economy. Luttmer (1996) investigates asset pricing in economies with frictions and shows that with transaction costs (bid-ask spreads) there is in general little evidence against the consumption-based power utility model with low risk-aversion parameters. Holmström and Tirole (1998) introduce a link between the liquidity demand of financially constrained firms and asset prices. Acharya and Pedersen (2004) propose a liquidity-adjusted CAPM model where transaction costs are time-varying. Alternatively, Vayanos

(2004) takes transactions costs as fixed but introduces the risk of having to liquidate a portfolio. Lagos (2006) extends the search friction argument to multiple assets: in a decentralized exchange, agents with uncertain future hedging demand prefer assets with lower search costs.

The rest of the paper is organized as follows. The next section presents the model and its state-space representation. Section II describes the data and Section III introduces the estimation method based on the UKF. We report estimation results for models with and without liquidity in Section IV. Section V evaluates the information content of liquidity for excess returns and interest rate spreads while Section VI identifies economic determinants of liquidity. Section VIII concludes.

I A Term Structure Model With Liquidity

We base our model on the Arbitrage-Free Extended Nelson-Siegel [AFENS] model introduced in CDR. This model belongs to the affine family (Duffie and Kan (1996)). The latent state variables relevant for the evolution of interest rates are grouped within a vector F_t of dimension $k = 3$. Its dynamics under the risk-neutral measure Q is described by the stochastic differential equation

$$dF_t^Q = K^Q(\theta^Q - F_t) + \Sigma dW_t^Q, \tag{1.1}$$

where dW_t is a standard Brownian motion process. Combined with the assumption that the short rate is affine in all three factors, the model then leads to the usual affine solution for discount bond yields.

In this context, CDR show that if the short rate is defined as $r_t = F_{1,t} + F_{2,t}$ and if the mean-reversion matrix K^Q is restricted to

$$K^Q = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix}, \tag{1.2}$$

then the absence of arbitrage opportunity implies the discount yield function,

$$y(F_t, m) = a(m) + F_{1,t}b_1(m) + F_{2,t}b_2(m) + F_{3,t}b_3(m), \tag{1.3}$$

with loadings given by

$$\begin{aligned} b_1(m) &= 1, \\ b_2(m) &= \left(\frac{1-\exp(-m\lambda)}{m\lambda}\right), \\ b_3(m) &= \left(\frac{1-\exp(-m\lambda)}{m\lambda} - \exp(-m\lambda)\right), \end{aligned} \tag{1.4}$$

where $m \geq 0$ is the length of time until maturity (see Appendix C for the $a(m)$ term).

These loadings are consistent with the static Nelson-Siegel representation of forward rates (Nelson and Siegel (1987), NS hereafter). Their shapes across maturities lead to the usual interpretations of factors in terms of level, slope and curvature. Moreover, the NS representation is parsimonious and imposes a smooth shape to the forward rate curve. Empirically, this approach is robust to over-fitting and delivers performance in line with, or better than, other methods for pricing out-of-sample bonds in the cross-section of maturities[9]. Conversely, its smooth shape is useful to identify deviations of observed yields from an idealized curve.

A dynamic extension of the NS model, the Extended Nelson-Siegel model [ENS], was first proposed by Diebold and Li (2006) and Diebold et al. (2006). Diebold and Li (2006) document large improvements in long-horizon interest rate forecasting. They argue that the ENS model performs better than the best essentially affine model of Duffee (2002) and point toward the model's parsimony to explain its successes. A persistent concern, though, was that the ENS model does not enforce the absence of arbitrage. This is precisely the contribution of CDR. They derive the class of continuous-time arbitrage-free affine dynamic term structure models with loadings that correspond to the NS representation. Intuitively, an AFENS model corresponds to a canonical affine model in Dai and Singleton (2000) where the loading shapes have been restricted through over-identifying assumptions on the parameters governing the risk-neutral dynamics of latent factors. CDR compare the ENS and AFENS models and show that implementing these restrictions improves forecasting performances further.

Interestingly, CDR show that we are free to choose the drift and variance

[9] See Bliss (1997) and Anderson et al. (1996) for an evaluation of yield curve estimation methods.

term for the dynamics under the physical measure

$$dF_t^P = K^P(\theta^P - F_t) + \Sigma dW_t^P, \tag{1.5}$$

and we impose that Σ is lower triangular and that K^P is diagonal[10]. We can then cast the model within a discretized state-space representation. The state equation becomes

$$(F_t - \bar{F}) = \Phi(F_{t-1} - \bar{F}) + \Gamma\epsilon_t, \tag{1.6}$$

where the innovation ϵ_t is standard Gaussian, the autoregressive matrix Φ is

$$\Phi = \exp\left(-K\frac{1}{12}\right) \tag{1.7}$$

and the covariance matrix Γ can computed from

$$\Gamma = \int_0^{\frac{1}{12}} e^{-Ks}\Sigma\Sigma^T e^{-Ks} ds. \tag{1.8}$$

Finally, we define a new latent state variable, L_t, that will be driving the liquidity premium. Its transition equation is

$$(L_t - \bar{L}) = \phi^l(L_{t-1} - \bar{L}) + \sigma^l\epsilon_t^l, \tag{1.9}$$

where the innovation ϵ_t^l is standard Gaussian and uncorrelated with ϵ_t.

Typically, term structure models are not estimated from observed prices. Rather, coupon bond prices are converted to forward rates using the bootstrap method. This is convenient as affine term structure models deliver forward rates that are linear in state variables. Is is also thought to be innocuous because bootstrapped forward rates achieve near-exact pricing of the original sample of bonds. Unfortunately, this extreme fit means that a naive application of the bootstrap pushes any liquidity effects and other price idiosyncracies into forward rates. Fama and Bliss (1987) handle this sensitivity to over-fitting by excluding bonds with "large" price differences relative to their neighbors.[11] This approach is certainly justified for many

??

[10]Formally, the assumption on Σ is required for identification purposes. In practice, the presence of the off-diagonal elements in the K^P matrix does not change our results. Moreover, CDR show that allowing for an unrestricted matrix K^P deteriorates out-of-sample performance.

[11]The CRSP data set of zero coupon yields is based on the approach proposed by Fama and Bliss. See also the CRSP documentation for a description of this procedure. Briefly, a first filter includes a quote if its yield to maturity falls within a range of 20 basis points

of the questions addressed in the literature, but it removes any evidence of large liquidity effects. Moreover, the FB data set focuses on discount bond prices at annual maturity intervals. This smooths away evidence of small liquidity effects remaining in the data and passed through to forward rates. These effects would be apparent from reversals in the forward rate function at short maturity interval. Consider three quotes for bonds with successive maturities $M_1 < M_2 < M_3$. A relatively expensive quote at maturity M_2 induces a relatively small forward rate from M_1 to M_2. However, the following normal quote with maturity M_3 requires a relatively large forward rate from M_2 to M_3. This is needed to compensate the previous low rate and to achieve exact pricing as required by the bootstrap. However, the reversal cancels itself as we sum intra-period forward rates to compute annual rates.

Instead of using smoothed data, we proceed from observed coupon bonds with maturity, say, M and with coupons at maturities $m = m_1, \ldots, M$. The price, $D_t(m)$, of a discount bond with maturity m, used to price intermediate payoffs, is given by

$$D_t(m) = \exp\left(-m(a(m) + b(m)^T F_t)\right) \qquad m \geq 0,$$

which follows directly from equation (1.3) but where we use vector notation for factors F_t and factor loadings $b(m)$. In a frictionless economy, the absence of arbitrage implies that the price of a coupon bond equals the sum of discounted coupons and principal. That is, the frictionless price is

$$P^*(F_t, Z_t) = \sum_{m=m_1}^{M} D_t(m) \times C_t(m), \tag{1.10}$$

where Z_t includes (deterministic) characteristics relevant for pricing a bond. In this case, it includes the maturity M and the schedule of future coupons and principal payments, $C_t(m)$.

However, with a short-sale constraint on government bonds and a collateral constraint in the repo market, Luttmer (1996) shows that the set of stochastic discount factors consistent with the absence of arbitrage satisfies

from one of the moving averages on the 3 longer or the 3 shorter maturity instruments *or* if its yield to maturity falls between the two moving averages. When computing averages, precedence is given to bills when available and this is explicitly designed to exclude the impact of liquidity on notes and bonds with maturity of less than one year. Amihud and Mendelson (1991) document that yield differences between notes and adjacent bills is 43 basis point on average, a figure much larger than the 20 basis point cutoff. The second filter excludes observations that cause reversals of 20 basis points in the bootstrapped discount yield function. The impact of these filters has not been studied in the literature.

$P \geq P^*$. These constraints match the institutional features of the Treasury market. An investor cannot issue new bonds to establish a short position. Instead, she must borrow the bond on the repo market through a collateralized loan. Then, we model the price, $P(F_t, L_t, Z_t)$, of a coupon bond with characteristics Z_t as the sum of discounted coupons to which we add a liquidity term,

$$P(F_t, L_t, Z_{n,t}) = \sum_{m=1}^{M_n} D_t(m) \times C_{n,t}(m) + \zeta(L_t, Z_{n,t}).$$

Here Z_t also includes the age of the bond. Note that the liquidity term should be positive to be consistent with a Luttmer (1996).

That the on-the-run premium is related to the short-sale constraint on government bonds and the collateral constraint in the repo market is justified by the results of Vayanos and Weill (2006) (see also Duffie (1996)). They show that the combination of these constraints with search frictions on the repo market induces differences in funding costs that favor recently issued bonds. Intuitively, the repo market provides the required heterogeneity between assets with identical payoffs. An investor cannot choose which bond to deliver to unwind a repo position; she must find and deliver the same security she had originally borrowed. Because of search frictions, then, investors are better off in the aggregate if they coordinate around one security to reduce search costs. In practice, the repo rate is lower for this special issue to provide an incentive for bond holders to bring their bonds to the repo market. Typically, recently issued bonds benefit from these lower financing costs, leading to the on-the-run premium. Moreover, these bonds offer lower transaction costs adding to the wedge between asset prices (Amihud and Mendelson (1986)). Empirically, both channels seem to be at work although the effect of lower transaction costs appears weaker than the effect of lower funding rates.[12]

Grouping observations together, and adding an error term, we obtain our measurement equation

$$P(F_t, L_t, Z_t) = C_t D_t + \zeta(L_t, Z_t) + \Omega \nu_t, \tag{1.11}$$

where C_t is the $(N \times M_{max})$ payoffs matrix obtained from stacking the N

[12] Amihud and Mendelson (1991) and Goldreich et al. (2005) consider transaction costs. Jordan and Jordan (1997), Krishnamurthy (2002) and Cheria et al. (2004) consider funding costs. See also, Buraschi and Menini (2002) for the German bonds market.

row vectors of individual bond payoffs and M_{max} is the longest maturity group in the sample. Shorter payoff vectors are completed with zeros. Similarly, $\zeta(L_t, Z_t)$ is a $N \times 1$ vector obtained by staking the individual liquidity premium. D_t is a $(M_{max} \times 1)$ vector of discount bond prices and the measurement error, ν_t, is a $(N \times 1)$ gaussian white noise uncorrelated with innovations in state variables. The matrix Ω is assumed diagonal and its elements are a linear function of maturity,

$$\omega_n = \omega_0 + \omega_1 M_n,$$

which reduce substantially the dimension of the estimation problem. However, leaving the diagonal elements of Ω unrestricted does not affect our results[13].

Our specification of the liquidity premium is based on a latent factor common to all bonds but with loadings that vary with maturity and age. The premium is given by

$$\zeta(L_t, Z_{n,t}) = L_t \times \beta_{M_n} \exp\left(-\frac{1}{\kappa}\ age_{n,t}\right) \tag{1.12}$$

where age_t is the age, in years, of the bond at time t. The parameter β_M controls the average on-the-run premium at each fixed maturity M. Warga (1992) document the impact of age and maturity on the average premium. We estimate β for a fixed set of maturities and the shape of β is unrestricted between these maturities.[14] Next, the parameter κ controls the on-the-run premium's decay with age. The gradual decay of the premium with age has been documented by Goldreich et al. (2005). For instance, immediately following its issuance (i.e.: $age = 0$), the loading on the liquidity factor is $\beta_M \times 1$. Taking $\kappa = 0.5$, the loading decreases by half within any maturity group after a little more than 4 months following issuance : $\zeta(L_t, 4) \approx \frac{1}{2}\zeta(L_t, 0))$. While the specification above reflects our priors about the impact of age and maturity, the scale parameters are left unrestricted at estimation and we allow for a continuum of shapes for the decay of liquidity. However, we fix $\beta_{10} = 1$ to identify the level of the liquidity factor with the average premium of a just-issued 10-year bond relative to a very old bond

[13]This may be due to the fact that the level factor explains most of yields variability. Its impact on bond prices is linear in duration and duration is approximately linear in maturity, at least for maturities up to 10 years. Bid-ask spreads increase with maturity and may also contribute to an increase in measurement errors with maturity.

[14]Opportunities of arbitrage may arise if β follows a step process across maturities. We thank an anonymous referee for this remark.

with the same maturity and coupons.

Equation (1.11) shows that omitting the liquidity term will push the impact of liquidity into pricing errors, possibly leading to biased estimators and large filtering errors. Alternatively, adding a liquidity term amounts to filtering a latent factor present in pricing errors. However, Equation (1.11) shows that this factor captures that part of pricing errors correlated with bond ages. Our maintained hypothesis is that any such positive factor can be interpreted as a liquidity effect. Clearly, the impact of age on the price of a bond can hardly be rationalized in a frictionless economy.

Intuitively, our specification delivers a discount rate function consistent with off-the-run valuation but remains silent on the linkage with the equilibrium stochastic discount factor. A structural specification of the liquidity premium raises important challenges. The on-the-run premium is a real arbitrage opportunity unless we explicitly consider the costs of shorting the more expensive bond or, alternatively, the benefits accruing to the bondholder from a lower repo rate. These features are absent from the current crop of term structure models with the notable exception of Cheria et al. (2004) who allow for a convenience yield, due to lower repo rates accruing to holders of an on-the-run issue. Clearly, theory suggests that using repo rates may improve the identification of the premium. Unfortunately, this would restrict our analysis to a much shorter sample where repo data are readily available. In any case, a joint model of the term structure of repo rates and of government yields may still not be free of arbitrage unless we also model the convenience yield of holding short-term government securities. This follows from the observation that a Treasury bill typically offers a lower yield than a repo contract with the same maturity. Moreover, the stochastic properties of repo rates are not well known, as well as the form of their relationships with bond yields. This is beyond the scope of this paper. Our strategy bypasses these challenging considerations but still uncovers the key role funding liquidity. We now turn to a description of the data.

II Data

We use end-of-month prices of U.S. Treasury securities from the CRSP data set. Our sample covers the period from January 1986 to December 2008. However, we estimate the model both with and without 2008 data. Before 1986, interest income had a favorable tax treatment compared to capital gains and investors favored high-coupon bonds. The resulting tax premium and the on-the-run premium cannot be disentangled in the earlier period. When interest rates are rising, recently issued bonds had relatively high coupons and were priced at a premium both for their liquidity and for their tax benefits. Green and Ødegaard (1997) document that the high-coupon tax premium mostly disappeared when the asymmetric treatment of interest income and capital gains was eliminated following the 1986 tax reform.

The CRSP data set[15] provides quotes on all outstanding U.S. Treasury securities. We filter unreliable observations and construct bins around maturities of 3, 6, 9, 12, 18, 24, 36, 48, 60, 84 and 120 months.[16] Then, at each date, and for each bin, we choose a pair of securities to identify the on-the-run premium. First, we want to pick the on-the-run security if any is available. Unfortunately, on-the-run bonds are not directly identified in the CRSP database. Instead, we use time since issuance as a proxy and pick the most recently issued security in each maturity bin. Second, we choose the security that most closely matches the bin's maturity (e.g. 3 months, 6 months,...). Note that pinning off-the-run securities at fixed maturities ensures a stable coverage of the term structure of interest rates. Also, by construction, securities within each pair have the same credit quality and very close times to maturity. We do not match coupon rates but coupon differences within pairs are low in practice.

The most important aspect of our sample is that whenever a security trades at premium relative to its pair companion, any large price difference cannot be rationalized from small coupon or maturity differences under the no-arbitrage restriction. On the other hand, price differences common across maturities and correlated with age will be attributed to liquidity. Note that the most recent issue for a given bin and date is not always an on-the-run security. This may be due to the absence of new issuance in some maturity bins throughout the whole sample (e.g. 18 months to maturity) or within some sub-periods (e.g. 84 months to maturity). Alternatively, the on-the-run bond may be a few months old, due to the quarterly issuance pattern observed in some maturity categories. In any case, this introduces variability in age differences which, in turn, identifies how the liquidity premium varies with age.

We now investigate some features of our sample of $265 \times 22 = 5830$ observations. The first two columns of Table I present means and standard deviations of age for each liquidity-maturity category. The average off-the-run

[15]See Elton and Green (1998) and Piazzesi (2005) for discussion of the CRSP data set.
[16]See Apppendix A for more details on data filter.

security is always older than the corresponding on-the-run security. Typically, the off-the-run security has been in circulation for more than a year. In contrast, the on-the-run security is typically a few months old and only a few weeks old in the 6 and 24-month categories. A relatively low average age for the recent issues indicates a regular issuance pattern. On the other hand, the relatively high standard deviations in the 36 and 84-month categories reflect the decision by the U.S. Treasury to stop the issuance cycles at these maturities.

[Table I about here.]

Next, Table I presents means and standard deviations of duration[17]. Average duration is almost linear in maturity. As expected, duration is similar within pairs implying that averages of cash flow maturities are very close. Finally, the last columns of Table I show that the term structure of coupons is upward sloping on average and the high standard deviations indicate important variations across the sample. This is in part due to the general decline of interest rates. Nonetheless, coupon rate *differences* within pairs are small on average. To summarize our strategy, differences in duration and coupon rates are kept small within each pair but differences of ages are highlighted so that we can identify any effect of liquidity on prices that is linked to age.

III Estimation Methodology

Equations (1.6), (1.9) and (1.11) can be summarized as a state-space system

$$\begin{aligned}(X_t - \bar{X}) &= \Phi_X(X_{t-1} - \bar{X}) + \Sigma_X \epsilon_t \\ P_t &= \Psi(X_t, C_t, Z_t) + \Omega \nu_t,\end{aligned} \tag{1.13}$$

where $X_t \equiv [F_t^T \; L_t]^T$ and Ψ is the (non-linear) mapping of cash flows C_t, bond characteristics, Z_t, and current states, X_t, into prices, P_t.

Estimation of this system is challenging because we do not know the joint density of factors and prices. Various strategies to deal with non-linear state-space systems have been proposed in the filtering literature: the Extended Kalman Filter (EKF), the Particle Filter (PF) and more recently the Unscented Kalman Filter[18] (UKF). The UKF is based on a method

[17] Duration is the relevant measure to compare maturities of bonds with different coupons.

[18] See Julier et al. (1995), Julier and Uhlmann (1996) and Wan and der Merwe (2001) for a textbook treatment. Another popular approach bypasses filtering altogether. It assumes that some prices are observed without errors and obtains factors by inverting the pricing equation. In our context, the choice of maturities and liquidity types that are not affected

for calculating statistics of a random variable which undergoes a nonlinear transformation. It starts with a well-chosen set of points with given sample mean and covariance. The nonlinear function is then applied to each point and moments are computed from transformed points. This approach has a Monte Carlo flavor but the sample is drawn according to a specific deterministic algorithm. It delivers second-order accuracy with no increase in computing costs relative to the EKF. Moreover, analytical derivatives are not required. The UKF has been introduced in the term structure literature by Leippold and Wu (2003) and in the foreign exchange literature by Bakshi et al. (2005). Recently, Christoffersen et al. (2007) compared the EKF and the UKF for the estimation of term structure models. They conclude that the UKF improves filtering results and substantially reduces estimation bias.

To set up notation, we state the standard Kalman filter algorithm as applied to our model. We then explain how the unscented approximation helps overcome the challenge posed by a non-linear state-space system. First, consider the case where Ψ is linear in X and where state variables and bond prices are jointly Gaussian. In this case, the Kalman recursion provides optimal estimates of current state variables given past and current prices. The recursion works off estimates of state variables and their associated MSE from the previous step,

$$\begin{aligned} \hat{X}_{t+1|t} &\equiv E\left[X_{t+1}|\Im_t\right], \\ Q_{t+1|t} &\equiv E\left[(\hat{X}_{t+1|t} - X_{t+1})(\hat{X}_{t+1|t} - X_{t+1})^T\right], \end{aligned} \tag{1.14}$$

where $\Im_t$ belongs to the natural filtration generated by bond prices. The associated predicted bond prices, and MSE, are given by

$$\begin{aligned} \hat{P}_{t+1|t} &\equiv E\left[P_{t+1}|\Im_t\right] \\ &= \Psi(\hat{X}_{t+1|t}, C_{t+1}, Z_{t+1}), \end{aligned} \tag{1.15}$$

$$\begin{aligned} R_{t+1|t} &\equiv E\left[(\hat{P}_{t+1|t} - P_{t+1})(\hat{P}_{t+1|t} - P_{t+1})^T\right] \\ &= \Psi(\hat{X}_{t+1|t}, C_{t+1}, Z_{t+1})^T \hat{Q}_{t+1|t} \Psi(\hat{X}_{t+1|t}, C_{t+1}, Z_{t+1}) + \Omega, \end{aligned} \tag{1.16}$$

using the linearity of Ψ. The next step compares predicted to observed bond

by measurement errors is not innocuous and impacts estimates of the liquidity factor.

prices and update state variables and their MSE,

$$\hat{X}_{t+1|t+1} = \hat{X}_{t+1|t} + K_{t+1}(P_{t+1} - \hat{P}_{t+1|t}), \tag{1.17}$$

$$Q_{t+1|t+1} = Q_{t+1|t} + K_{t+1}^T (R_{t+1|t})^{-1} K_{t+1}, \tag{1.18}$$

where

$$\begin{aligned} K_{t+1} &\equiv E\left[(\hat{X}_{t+1|t} - X_{t+1})(\hat{P}_{t+1|t} - P_{t+1})^T\right], \\ &= Q_{t+1|t}\Psi(\hat{X}_{t+1|t}, C_{t+1}, Z_{t+1}), \end{aligned} \tag{1.19}$$

measures co-movements between pricing and filtering errors. Finally, the transition equation gives us a conditional forecast of X_{t+2},

$$\hat{X}_{t+2|t+1} = \Phi_X \hat{X}_{t+1|t+1}, \tag{1.20}$$

$$Q_{t+2|t+1} = \Phi_X^T Q_{t+1|t+1} \Phi_X + \Sigma_X \Sigma_X^T. \tag{1.21}$$

The recursion delivers series $\hat{P}_{t|t-1}$ and $R_{t|t-1}$ for $t = 1, \cdots, T$. Treating $\hat{X}_{1|0}$ as a parameter, and setting $R_{1|0}$ equal to the unconditional variance of measurement errors, the sample log-likelihood is

$$L(\theta) = \sum_{t=1}^{T} l(P_t; \theta) = \sum_{t=1}^{T} \left[\log \phi(\hat{P}_{t+1|t}, R_{t+1|t})\right], \tag{1.22}$$

where $\phi(\cdot, \cdot)$ is the multivariate Gaussian density.

However, because $\Psi(\cdot)$ is not linear, equations (1.15) and (1.16) do not correspond to the conditional expectation of prices and the associated MSE. Also, (1.19) does not correspond to the conditional covariance between pricing and filtering errors. Still, the updating equations (1.17) and (1.18) remain justified as optimal linear projections. Then, we can recover the Kalman recursion provided we obtain approximations of the relevant conditional moments. This is precisely what the unscented transformation achieves, using a small deterministic sample from the conditional distribution of factors while maintaining a higher order approximation than linearization[19]. We can then use the likelihood given in (1.22), but in a QML context. Using standard results, we have $\hat{\theta} \approx N(\theta_0, T^{-1}\Omega)$ where $\hat{\theta}$ is the QML estimator of θ_0 and the covariance matrix is

$$\Omega = E\left[\left(\zeta_H \zeta_{OP}^{-1} \zeta_H\right)^{-1}\right], \tag{1.23}$$

[19] See Appendix B.

where ζ_H and ζ_{OP} are the alternative representations of the information matrix, in the Gaussian case. These can be consistently estimated via their sample counterparts. We have

$$\hat{\zeta_H} = -T^{-1}\left[\frac{\partial^2 L(\hat{\theta})}{\partial\theta\partial\theta'}\right] \tag{1.24}$$

and

$$\hat{\zeta_{OP}} = T^{-1}\sum_{t=1}^{T}\left[\left(\frac{\partial l(t,\hat{\theta})}{\partial\theta}\right)\left(\frac{\partial l(t,\hat{\theta})}{\partial\theta}\right)^T\right]. \tag{1.25}$$

Finally, the model implies some restrictions on the parameter space. In particular, ϕ_l and diagonal elements of Φ must lie in $(-1,1)$ while κ and λ must remain positive. In practice, large values of κ or λ lead to numerical difficulties and are excluded. Finally, we maintain the second covariance contour of state variables inside the parameter space associated with positive interest rates. The filtering algorithm often fails outside this parameter space. None of these constraints binds around the optimum and estimates remain unchanged when the constraints are relaxed.

IV Estimation Results

We first estimate a restricted version of our model, excluding liquidity. Filtered factors and parameter estimates are consistent with results obtained by CDR from zero-coupon bonds. More interestingly, the on-the-run premium reveals itself in the residuals from the benchmark model. This provides a direct justification for linking the premium with the age and maturity of each bond. We then estimate the unrestricted liquidity model. The null of no liquidity is easily rejected and the liquidity factor captures systematic differences between on-the-run and off-the-run bonds. Finally, estimates imply that the on-the-run premium increases with maturity but decreases with the age of a bond.

A Estimation Without Liquidity

Estimation[20] of the benchmark model put the curvature parameter at $\hat{\lambda} = 0.6786$ when time periods are measured in years. The standard error is 0.0305 and 0.0044 when using the QMLE and MLE covariance matrix, respectively.

[20] Estimation is implemented in MATLAB via the *fmincon* routine with the medium-scale (active-set) algorithm. Different starting values were used. For standard errors computations, we obtain the final Hessian update (BFGS formula) and each observation gradient is obtained through a centered finite difference approximation evaluated at the optimum.

This estimate pins the maximum curvature loading at a maturity close to 30 months.

[Table II about here.]

Figure 1.1 displays the time series of the liquidity (Panel (a)) and the term structure (Panel (b)) factors. Estimates for the transition equation are given in Table IIa. The results imply average short and long term discount rates of 3.73% and 5.45%, respectively. The level factor is very persistent, perhaps a unit root. This standard result in part reflects the gradual decline of interest rates in our sample. The slope factor is slightly less persistent and exhibits the usual association with business cycles. Its sign changes before the recessions of 1990 and 2001. The slope of the term structure is also inverted starting in 2006, during the so-called "conundrum" episode. Finally, the curvature factor is closely related to the slope factor.

Standard deviations of pricing errors are given by

$$\begin{aligned}\sigma(M_n) = \quad & 0.0229 \quad + \quad 0.0284 \times M_n, \\ & (0.017,\ 0.0012) \quad (0.021,\ 0.0006)\end{aligned}$$

with QMLE and MLE standard errors for each parameter. This implies standard deviations of %0.05 and \$0.31 dollars for maturities of 1 and 10 years, respectively. Using durations of 1 and 7 years, this translates into yield errors of 5.1 and 4.4 bps. Table IIIa gives more information on the fit of the benchmark model. Root Mean Squared Errors (RMSE) increase from \$0.047 and \$0.046 for 3-month on-the-run and off-the-run securities, respectively, to \$0.35 and \$0.39 at 10-year maturity. As discussed above, the monotonous increase of RMSE with maturity reflects the higher sensitivity of longer maturity bonds to interest rates. It may also be due to higher uncertainty surrounding the true prices, as signaled by wider bid-ask spreads. In addition, for most maturities, the RMSE is larger for on-the-run bonds. For the entire sample, the RMSE is \$0.188.

Notwithstanding differences between estimation approaches, our results are consistent with CDR. Estimating using coupon bonds or using bootstrapped data provides similar pictures of the underlying term structure of interest rates. Also, the approximation introduced when dealing with nonlinearities is innocuous. However, preliminary estimation of forward rate curves smooths away any effect of liquidity. In contrast, our sample comprises on-the-run and off-the-run bonds. Any systematic price differences not due to cash flow differences will be revealed in the pricing errors.

[Table III about here.]

Table IIIa confirms that Mean Pricing Errors (MPE) are systematically higher for on-the-run securities. On-the-run residuals are systematically higher than off-the-run residuals. For a recent 12-month T-Bill, the average difference is close to $0.08, controlling for cash flow differences. Similarly, a recently issued 5-year bond is $0.25 more expensive on average than a similar but older issue.[21] To get a clearer picture of the link between age and price differences, consider Figure 1.4. The top panels plot residual differences within the 12-month and 48-month categories. The bottom panels plot the ages of each bond in these categories. Panel (c) shows that the U.S. Treasury stopped regular issuance of the 12-month Notes in 2000. The liquidity premium was generally positive until then but stopped when issuance ceased. Afterwards, each pair is made of old 2-year Notes, and evidence of a premium disappears from the residuals. Panel (d) shows that there has been regular issuance of 4-year bonds early in the sample. As expected, the difference between residuals is generally positive whenever there is a significant age difference between the two issues. Moreover, in each case, on-the-run (i.e. low age) bonds appear overpriced compared to off-the-run (i.e. high age) bonds. This correspondence between issuance patterns and systematic pricing errors can be observed in each maturity category. The premium increases with maturity but decreases with age.

Bonds with 24 months to maturity seem to carry a smaller liquidity premium than what would be expected given the regular monthly issuance for this category. Note that a formal test rejects the null hypothesis of zero-mean residual differences. Interestingly, Jordan and Jordan (1997) could not find evidence of a liquidity or specialness effect at that maturity[22]. A smaller price premium for 2-year Notes is intriguing and we can only conjecture as to its causes. Recall that the magnitude of the premium depends on the benefits of higher liquidity, both in terms of lower transaction costs and lower repo rates. However, it also depends on the expected length of time a bond will offer these benefits. Results in Jordan and Jordan (1997) suggests that 2-year Notes remain "special" for shorter periods of time (see Table I, p.2057). Similarly, Goldreich et al. (2005) find that the on-the-run premium on 2-year Notes goes to zero faster than other maturities, on average. This is

[21] Note that the price impact of liquidity increases with maturity. This is consistent with the results of Amihud and Mendelson (1991).

[22] See Jordan and Jordan (1997) p. 2061: "With the exception of the 2-year notes [...], the average price differences in Table II are noticeably larger when the issue examined is on special."

consistent with its short issuance cycle. Alternatively, holders of long-term bonds may re-allocate funds from their now short maturity bonds into newly issued longer term securities. If the two-year mark serves as a focus point for buyers and sellers, this may cause a larger volume of transactions around this key maturity, increasing the liquidity value of surrounding assets.

B Estimation With Liquidity

Estimation of the unrestricted model leads to a substantial increase of the log-likelihood. The benchmark model is nested with 15 parameter restrictions and the improvement in likelihood is such that the LR test-statistic leads to a p-value that is essentially zero[23]. The estimate for the curvature parameter is now $\hat{\lambda} = 0.7304$ with QMLE and MLE standard errors of 0.0857 and 0.0043. Results for the transition equations are given in Table IIb. These imply average short and long term discount rates of 4.09% and 5.76% respectively. Interestingly, the yield curve level is higher once we account for the liquidity premium. Intuitively, the off-the-run yield curve is higher than an otherwise unadjusted estimate would suggest. The standard deviations of measurement errors are given by

$$\begin{array}{rcc} \sigma(M_n)^2 = & 0.0227+ & 0.0251 \times M_n, \\ & (0.016, 0.001) & (0.0021, 0.0006) \end{array}$$

with QMLE and MLE standard errors for each parameter in parenthesis. Then, standard deviations are \$0.048 and \$0.274 for bonds with one and ten years to maturity, respectively. Using durations of 1 and 7, this translates into standard deviations of 4.8 and 3.9 bps when measured in yields. Overall, parameter estimates and latent factors are relatively unchanged compared to the benchmark model.

We estimate the decay parameter at $\hat{\kappa} = 1.89$ with QMLE and MLE standard errors of 1.23 and 0.45 respectively. Estimates of β are given in Table IV. Note that the level of the liquidity premium increases with maturity.[24] The pattern accords with the observations made from residuals of the model without liquidity. Moreover, Table IIIa shows that the model eliminates most of the systematic differences between on-the-run and off-the-run bonds. There is still some evidence of a systematic difference in the 10-year category where the average error decreases from \$0.31 to \$0.26. We conclude

[23] The benchmark model reached a maximum at 1998.6 while the liquidity model reached a maximum at 3482.6.

[24] The estimated average level is lower in the 10-year group relative to the 5-year and 7-year group. This is due to the lower average age of bonds in this groups.

that part of the variations in the 10-year on-the-run premium is not common with variations in other maturity groups. Finally, Table IIIb shows RMSE improvements for almost all maturities while the overall sample RMSE decreases from $0.188 to $0.151.

[Table IV about here.]

Figure 1.5 draw the residual differences within the 12-month and 60-month category, respectively. This is another way to see that the model removes systematic differences between residuals. Overall, the evidence points toward a large common factor driving the liquidity premium of on-the-run U.S. Treasury securities. We interpret this liquidity factor as a measure of the value of funding liquidity to investors. The results below show that its variations also explain a substantial share of the risk premia observed in different interest rate markets.

V Liquidity And Bond Risk Premia

In this section, we present evidence that variations in the value of funding liquidity, as measured from a cross-section of on-the-run premia, share a common components with variations of risk premia in other interest rate markets. In other words, conditions prevailing on the funding market induce an aggregate risk factor that affects each of these markets. Of course, an increase in the liquidity factor necessarily leads to lower excess returns for on-the-run bonds. We show here that it also leads to lower risk premia for off-the-run bonds as well as higher risk premia on LIBOR loans, swap contracts and corporate bonds. Thus, although the payoffs of these assets are not directly related to the higher liquidity of on-the-run securities, their risk premium and, hence, their price, is affected by a common liquidity factor. To summarize, exposure to liquidity risk in the U.S. Treasury funding market carries a substantial price of risk in the cross-section of bond returns. The impact across assets is similar to the often cited "flight-to-liquidity" phenomenon but remains pervasive in normal market conditions. This commonality across liquidity premia accords with a substantial theoretical literature supporting the existence of an economy-wide liquidity premium (Svensson (1985), Bansal and Coleman (1996), Holmtröm and Tirole (1998, 2001), Acharya and Pedersen (2004), Vayanos (2004), Lagos (2006), Brunnermeier and Pedersen (2008), Krishnamurthy and He (2008).). The following section presents our results.[25]

[25] All the results below are robust to choice of the off-the-run yield curve used to compute excess returns or spreads. Unless otherwise stated we use off-the-run yields

A Off-The-Run U.S. Treasury Bonds

We first document the negative relationship between liquidity and expected excess returns on off-the-run bonds. This is the return, over a given investment horizon, from holding a long maturity bond, in excess of the risk-free rate for that horizon. Figure 1.3a displays annual excess returns on a 2-year off-the-run bond along with the liquidity factor. The negative relationship is visually apparent throughout the sample but note the sharp variations around the crash of October 1987, the Mexican Peso crisis late in 1994, around the LTCM crisis in August 1998 and until the end of the millennium. At first, this tight link between on-the-run premia and returns from off-the-run Treasury bonds may be surprising. Recall that on-the-run bonds trade at a premium due to their anticipated transaction costs and funding advantages on the cash and repo markets. However, off-the-run bonds can be readily converted into cash via the repo market. This is especially true relative to other asset classes. In that sense, seasoned bonds are close substitutes to on-the-run bonds. Then, the risk premium of all Treasury bonds decreases in periods of high demand for the relative funding liquidity of on-the-run bonds. Longstaff (2004) documents price differences between off-the-run U.S Treasury bonds and Refcorp bonds[26] with similar cash flows. He argues that discounts on Refcorp bond are due to "...the liquidity of Treasury bonds, especially in unsettled markets.".

[Table V about here.]

We test this hypothesis through predictive regressions of off-the-run bond excess returns on the liquidity factor. We use the off-the-run curve from the model to compute excess returns and include term structure factors to control for the information content of forward rates (Fama and Bliss (1987), Campbell and Shiller (1991), Cochrane and Piazzesi (2005a)). The term structure factors spans forward rates but do not suffer from their near-collinearity. Table V presents the results. We consider (annualized) excess returns from holding off-the-run bonds with maturities of 2, 3, 4, 5, 7 and 10 years and for investment horizons of 1, 3, 6, 12, and 24 months. First, Panel (a) presents

from the Svensson, Nelson and Siegel method (Gurkaynak et al. (2006)) available at (http://www.federalreserve.gov/pubs/feds/2007). Using model-implied zero-coupon yields does not affect the results. Also, for ease of interpretation, we standardize each regressor by subtracting its mean and dividing by its standard deviation. For each risk premium regression, the constant corresponds to an estimate of the average risk premium and the coefficient on the liquidity factor measures the impact on expected returns, in basis points, of a one-standard deviation shock to liquidity.

[26] Refcorp is an agency of the U.S. government. Its liabilities have their principals backed with U.S. Treasury bonds and coupons explicitly guaranteed by the U.S. Treasury.

average risk premia. These range from 153 to 471 bps at one-month horizon and from 69 to 358 bps at annual horizon. These large excess returns are consistent with an average positive term structure slope and with a period of declining interest rates. Panel (b) presents estimates of the liquidity coefficients. The results are conclusive. Estimates are negative and significant at all horizons and maturities. Moreover, the impact of liquidity on excess returns is economically significant. At a one-month horizon a one-standard deviation shock to our measure of funding liquidity lowers expected excess returns obtained from off-the-run bonds by 187 and 571 bps for maturities of two and ten years respectively. At this horizon, R^2 statistics range from 7.34% to 4.23% (see Panel (c)). Regressions based on excess returns at an annual horizon correspond to the cased studied by Cochrane and Piazzesi (2005a) who document the substantial predictability of US Treasury excess returns from forward rates. The impact of funding liquidity is substantial. A one-standard deviation shock decreases expected excess returns by 103 basis points at 2-year maturity and by as much as 358 basis points at 10-year maturity. At this horizon, R^2 are substantially higher, ranging from 43% and 50%. Of course, these coefficients of variation pertain to the joint explanatory power of all regressors. Panel (c) also presents, in bracket, the R^2 of the same regressions but excluding the liquidity factor. The liquidity factor accounts for more or less half of the predictive power of the regressions.

The regressions above used excess returns and term structure factors computed from the term structure model. One concern is that model misspecification leads to estimates of term structure factors that do not correctly capture the information content of forward rates or that it induces spurious correlations between excess returns and liquidity. As a robustness check against both possibilities, we re-examine the predictability regressions but using excess returns and forward rates available from the CRSP zero-coupon yield data set. From this alternative data set, we compute annual excess returns on zero-coupon bonds with maturity from 2 to 5 years. As regressors, we include annual forward rates from CRSP at horizon from 1 to 5 years along with the liquidity factor from the model. Table VIa presents the results. Estimates of the liquidity coefficients are very close to our previous results (see Table Vb) and highly significant. We conclude that the predictability power of the liquidity factor is robust to how we compute excess returns and forward rates.

[Table VI about here.]

Furthermore, this alternative set of returns allows to check whether the

AFENS model captures important aspects of observed excess returns. Table VIb provides results for the regressions of CRSP excess returns on CRSP forward rates, excluding the liquidity factor. This is a replication of the unconstrained regressions in Cochrane and Piazzesi (2005a) but for our shorter sample period. This exercise confirms their stylized predictability results in this sample. That is, the predictive power of forward rates is substantial and we recover a tent-shaped pattern of coefficients across maturities. Next, Table VIc provides results of a similar regressions with CRSP forward rates but using excess returns computed from the *model*. Comparing the last two panels, we see that average excess returns, forward rate coefficients, as well as R^2s are similar across data sets. This is striking given that excess returns were recovered using very different approaches. The AFENS model captures the stylized facts of bond risk premia, which is an important measure of success for term structure models.[27]

The evidence shows that variations of funding liquidity value induce variations in the liquidity premium of Treasury bonds. Empirically, off-the-run US Treasury bonds are viewed as liquid substitutes to their recently issued counterparts and provide a hedge against fluctuations in funding liquidity. Note that this link between conditions on the funding market and the risk premium on a Treasury bond can hardly be attributed to traditional explanations of bond risk premia such as inflation risk or interest rate risk. Instead, we argue that frictions in the financial intermediation sector affect the Treasury market. The following section considers the impact of funding liquidity on LIBOR rates.

B LIBOR Loans

In this section, we link variations of the liquidity factor with variations in the risk compensation from money market loans. We consider the returns obtained from rolling over a lending position in the London inter-bank market at the LIBOR rate and funding this position at a fixed rate. This measures the reward of providing liquidity in the inter-bank market. In contrast with the government bond market, higher valuation of funding liquidity predicts higher excess returns. Figure 1.3b highlights the positive correlation between

[27] Fama (1984b) originally identified this modeling challenge but see also Dai and Singleton (2002). Other stylized facts are documented in Fama (1976), (1984a), and(1984b), as well as Startz (1982) for maturities below 1 year. See also Shiller (1979), Fama and Bliss (1987), Campbell and Shiller (1991). Our conclusions hold if we use Campbell and Shiller (1991) as a benchmark. We also conclude that the empirical facts highlighted by Cochrane and Piazzesi (2005a) are not an artefact of the bootstrap method. See the discussion in Dai et al. (2004) and Cochrane and Piazzesi (2005b).

liquidity and rolling excess returns. Again, note the spikes in 1987, 1994, in 1998 and around the end of the millennium.

Thus, interbank loans are poor substitutes to U.S. Treasury securities in time of funding stress. The reward for providing funds in the inter-bank market is higher when the relative value of on-the-run bonds increases. Thus, the spread of a LIBOR rate above the Treasury yield reflects the opportunity costs, in terms of future liquidity, of an interbank loan compared to the liquidity of a Treasury bonds on the repo or the cash markets. Indeed, in order to convert a loan back to cash, a bank must enter into a new bilateral contract to borrow money. The search costs of this transaction depend on the number of willing counterparties in the market and it may be difficult at critical times to convert a LIBOR position back to cash.[28]

As in the previous section, we test this hypothesis formally through predictive regressions of excess rolling returns on the liquidity factor. Again, we use term structure factors to control for the information content of forward rates. We consider investment horizons of 1, 3, 6, 12 and 24 months and rolling investments in LIBOR loans with 1, 3, 6 and 12 months to maturity. The LIBOR data is available from the web site of the BBA and we use a sample from January 1987 to December 2007. Table VII presents the results. For each loan maturity, the average excess returns is around 25 bps for the shortest horizon. Returns then decrease with longer horizon and become negative at the longest horizons. This reflects the average positive slope of the term structure. In practice, funding rolling short-term investments at a fixed rate does not produce positive returns on average. Still, the impact of liquidity is unambiguously positive for all horizons and maturities with t-statistics above 5 in most cases.

Interestingly, the impact of the liquidity *increases* with the horizon. A one-standard deviation shock to the value of liquidity increases returns on a rolling investment in one-month LIBOR loans by 16 and 90 bps at horizons 3 and 24 months, respectively. Results are similar for other maturities. In fact, the impact is sufficiently large that returns are positive on average, and the risk premium is higher than the slope of the term structure. This reflects the persistence of the liquidity premium. The R^2 from these regressions range from 30% to 50%. Moreover, the contribution of the liquidity factor to the predictability of LIBOR returns is substantial, generally doubling the R^2, or more. In the case of annual excess rolling returns from 3-month loans,

[28] Note that this does not preclude that part of the LIBOR spread is due to the higher default risk of the average issuer compared to the U.S. government.

the predictive power increases from 10.8% to 43.2% when we include the liquidity factor.

An alternative indicator of ex-ante returns from investment in the interbank market is the simple spread of LIBOR rates above risk-free zero-coupon yields. As an alternative test, we compute LIBOR spreads on loans with maturities of 1, 3, 6 and 12 months and consider regressions of these spreads on the liquidity and term structure factors. Panel (c) shows the positive relationship between liquidity and the 12-month LIBOR spread. Table VIIIa presents results from the regressions. A one-standard deviation shock to liquidity is associated with concurrent increases of 16, 12, 8 and 6 bps for loans with maturity of 1, 3, 6 and 12 months, respectively.

C Swap Spreads

The impact of funding liquidity extends to the swap market. This section documents the link between the liquidity factor and the spread of swap rates above the off-the-run curve. To the extent that swap rates are determined by anticipations of future LIBOR rates, results from the previous section suggest that swap spreads increase with the liquidity factor. Moreover, variations in funding liquidity may affect the swap market directly since the same intermediaries operate in the Treasury and the swap markets. We do not distinguish between these alternative channels here.

[Table VIII about here.]

We obtain a sample of swap rates from DataStream, starting in April 1987 and up to December 2007. We focus on swaps with maturities of 2, 5, 7 and 10 years and compute their spreads above the yield to maturity of the corresponding off-the-run par coupon bond. Figure 1.3d compares the liquidity factor with the 5-year swap spread. The positive relationship is apparent. Table VIIIb shows the results from regressions of swap spreads on the liquidity and term structure factors. First, the average spread rises with maturity, from 44 to 53 bps, and extends the pattern of LIBOR risk premia. Next, estimates of the liquidity coefficients imply that, controlling for term structure factors, a one-standard deviation shock to liquidity raises swap spreads from 5 to 7 basis points across maturities. The estimates are significant, both statically and economically, given the higher price sensitivities of swap to change in yields. For a 5-year swap with duration of 4.5, say, the price impact of a 6 basis point change is $0.27 while the price impact of the 6.3 bps rate change for a 1-year LIBOR loand is $0.063.[29] Finally, the

[29]We do not use returns on swap investment to measure expected returns. Swap investment requires zero initial investment. Determining the proper capital-at-risk to use

explanatory power of liquidity is high and increases with maturity.

Interestingly, funding liquidity affects swap spreads and LIBOR spreads similarly. This suggests that anticipations of liquidity compensation in the interbank loan market, rather than liquidity risk, is the main driver behind the aggregate liquidity component of swap risk premium. This supports previous literature (Grinblatt (2001), Duffie and Singleton (1997), Liu et al. (2006) and Fedlhütter and Lando (2007)) pointing toward LIBOR liquidity premium as an important driver of swap spreads. However, we show that the liquidity risk underlying a substantial part of that premium is not specific to the LIBOR market but reflects risks faced by intermediaries in funding markets.

D Corporate Spreads

The impact of funding liquidity extends to the corporate bond market. This section measures the impact of the liquidity factor on the risk premium offered by corporate bonds. Empirically, we find that the impact of liquidity has a "flight-to-quality" pattern across credit ratings. Following an increase of the liquidity factor, excess returns decrease for the higher ratings but increase for the lower ratings. Our results are consistent with the evidence that default risk cannot rationalize corporate spreads. Collin-Dufresne et al. (2001) find that most of the variations of non-default corporate spreads are driven by a single latent factor. We formally link this factor with funding risk. Our evidence is also consistent with the differential impact of liquidity across ratings found by Ericsson and Renault (2006). However, while they relate bond spreads to bond-specific measures of liquidity, we document the impact of an aggregate factor in the compensation for illiquidity.

Our analysis begins with Merrill Lynch corporate bond indices. We consider end-of-month data from December 1988 to December 2007 on 5 indices with credit ratings of AAA, AA, A, BBB and High Yield [HY] ratings (i.e. HY Master II index), respectively. In a complementary exercise, below, we use a sample of NAIC transaction data.[30] As in earlier sections, we measure the impact of liquidity on corporate bonds through predictive excess returns regressions. For each index, and each month, we compute returns in excess of the off-the-run zero coupon yield for investment horizons of 1, 3, 6, 12 and 24 months. We then project returns on the liquidity and term structure

in returns computation is somewhat arbitrary. It should be clear from Figure 1.3d that receiving fixed, and being exposed to short-term LIBOR fluctuations, will provide greater compensation when the liquidity premium is elevated.

[30]We thank Jan Ericsson for providing the NAIC transaction data and control variables. See Ericsson and Renault (2006) for a discussion of this data set.

factors. Again, term structure factors are included to control for the information content of the yield curve. The first Panel of Table IX presents the results.

First, as expected, average excess returns are higher for lower ratings. Next, estimates of the liquidity coefficients show that the impact of a rising liquidity factor is negative for the higher ratings and becomes positive for lower ratings. A one-standard deviation shock to the liquidity factor leads to decreases in excess returns for AAA, AA and A ratings but to increases in excess returns for BBB and HY ratings. Excess returns decrease by 2.27% for AAA index but increase by 2.38% for the HY index. For comparison, the impact on Treasury bonds with 7 and 10 years to maturity was -4.52% and -5.42%. Thus, on average, high quality bonds were considered substitutes, albeit imperfect, to U.S. Treasuries as a hedge against variations in funding conditions. On the other hand, lower-rated bonds were exposed to funding market shocks.

The differential impact of liquidity on excess returns across ratings suggests a flight-to-liquidity pattern. We consider an alternative sample, based on individual bond transaction data from the NAIC. While this sample covers a shorter period, from February 1996 until December 2001, the sample comprises actual transaction data and provides a better coverage of the rating spectrum. Once restricted to end-of-month observations, the sample includes 2,171 transactions over 71 months. To preserve parsimony, we group ratings in five categories.[31] We consider regression of NAIC corporate spreads on the liquidity and term structure factors but we also include the control variables used by Ericsson and Renault (2006). These are the VIX index, the returns on the S&P500 index, a measure of market-wide default risk premium and an on-the-run dummy signalling whether that particular bond was on-the-run at the time of the transaction. Control variables also include the level and the slope of the term structure of interest rates.[32].

The panel regressions of credit spreads for bond i at date t are given by

$$sprd_{i,t} = \alpha + \beta_1 L_t I(G_i = 1) + \cdots + \beta_5 L_t I(G_i = 5) + \gamma_h^T X_t + \epsilon_{i,t} \quad (1.26)$$

where L_t is the liquidity factor and $I(G_i = j)$ is an indicator function equal

[31] Group 1 includes ratings from AAA to A+, group 2 includes ratings A and A-, group 3 includes ratings BBB+, BBB and BBB-, group 4 includes ratings CCC+, CCC and CCC- while group 5 includes the remaining ratings down to C-

[32] We do not include individual bond fixed-effects as our sample is small relative to the number (998) of securities.

to one if the credit rating of bond i belongs in group $j = 1, \ldots, 5$. Control variables are grouped in the vector X_{t+h}. Table IXb presents the results. The flight-to-quality pattern clearly emerges from the results. For the highest rating category, an increase in liquidity value of one standard deviation decreases spreads by 31 and 20 basis points in groups 1 and 2 respectively. The effect is smaller and statistically undistinguishable from zero for group 3. Coefficients then become positive implying increases in spreads of 25 and 26 basis points for groups 4 and 5, respectively. This is an average effect through time and across ratings within each group.[33]

The results obtained from spreads computed from Merrill Lynch indices and spreads computed from NAIC transactions differ. While results from Merrill Lynch were inconclusive, estimates of liquidity coefficients obtained from NAIC data confirm that a shock to funding liquidity leads to lower corporate spreads in the highest rating groups but higher corporate spreads in the lowest rating groups. Two important differences between samples may explain the results. First, the composition of the index is different from the composition of NAIC transaction data. The impact of liquidity on corporate spreads may not be homogenous across issues. For example, the maturity or the age of a bond, the industry of the issuer and security-specific option features may introduce heterogeneity. Second, Merrill Lynch indices cover a much longer time span. The pattern of liquidity premia across the quality spectrum may be time-varying.

E Discussion

Focusing on the common component of on-the-run premia filters out local or idiosyncratic demand and supply effects on Treasury bond prices. The results above show that this measure of funding liquidity is an aggregate risk factor affecting money market instruments and fixed-income securities. These assets carry a significant, time-varying and common liquidity premium. That is, when the value of the most-easily funded collateral rises relative to other securities, we observe variations in risk premia for off-the-run U.S. government bonds, eurodollar loans, swap contracts, and corporate bonds. Empirically, the impact of aggregate liquidity on asset pricing appears strongly during crisis and the pattern is suggestive of a flight-to-quality behavior. Nevertheless, its impact is pervasive even in normal times.

[33]We do not report other coefficients. Briefly, the coefficient on the level factor is negative and significant. All other coefficients are insignificant but these results are are not directly comparable with Ericsson and Renault (2006) due to differences of models and sample frequencies.

Note that these regressions assumed a stable relationship between risk premium and funding liquidity. One important alternative is that the sign and the size of the impact of funding conditions itself depend on the intensity of the funding shock, as suggested by the recent experience. In particular, while corporate bonds with high ratings may be substitutes to Treasury bonds in good times, they experience large risk premium increases in funding crisis. We leave this for further research but note that this may explain the weak statistical evidence above in the case of corporate bonds. In any case, the main result of this section is that a substantial fraction of the risk premia is linked to variations in funding liquidity.

Jointly, the evidence is hard to reconcile with theories based on variations of default probability, inflation or interest rates and their associated risk premia. Instead, we link risk premium variations with conditions in the funding markets. This supports the theoretical literature that emphasizes the role of borrowing constraints faced by financial intermediaries (Gromb and Vayanos (2002), He and Krishnamurthy (2007)) and, in particular, that highlights the role of funding markets in financial intermediation (Brunnermeier and Pedersen (2008)). Different securities serve, in part, and to varying degrees, to fulfill investors' uncertain future needs for cash and their risk premium depend on the ability of intermediaries to provide immediacy in each market. In this context, it is interesting that the liquidity premium of government bonds appears to *decrease* when funding liquidity become scarce. This confers a special status to government obligations, and possibly to high-quality corporate bonds, as a hedge against variations in funding liquidity. We leave for further research the cause of this special attribute of government bonds. The next section identifies candidate determinants of liquidity valuation and characterizes aggregate liquidity in terms of known economic indicators.

VI Determinants Of Liquidity Value

The liquidity factor aggregates very diverse economic information. The value of liquidity services on the funding market depends on investors' demand for immediacy on markets where intermediaries are active. Next, funding costs will also vary with the capital position and the access to capital (present and future) of financial intermediaries that obtain leverage through secured loans. Finally, conditions on the funding market are affected by the availability of funds and, thus, by the relative tightness of monetary policy. In this section, we find that the value of funding liquidity, measured by the on-the-run factor, varies with changes in monetary aggregates and in bank reserves. Also, the value of funding liquidity increases with aggregate wealth and aggregate uncertainty as measured by valuation ratios and option-implied volatility of the SP500 stock index. Finally, the on-the-run premium rises when recently issued bonds offers relatively lower bid-ask spreads[34].

A Macroeconomic Variables

Ludvigson and Ng (2009) [LN hereafter] summarize 132 US macroeconomic series into 8 principal components. They then explore parsimoniously the predictive content of this large information set for bond returns. Their main result is that that a "real" and an "inflation" factor[35] have substantial predictive power for bond excess returns beyond the information content of forward rates. They also find that a "financial" factor is significant but that much of its information content is subsumed in the Cochrane-Piazzesi measure of bond risk premium.

Table X displays results from a regression of liquidity on macroeconomic factors (Regression A) from LN.[36] This shows that the funding liquidity factor shares tight linkages with the macroeconomy. Macroeconomic factors with significant coefficients are $F1$, $F2$ and $F4$, the "real", "financial" and "inflation" factors of LN that also predict bond risk premium. In addition, factors $F5$, $F6$, and $F7$ are also significant. As we discuss below, $F6$, and $F7$ can be interpreted as "monetary conditions" factors and $F5$ is a "housing activity" factor. Finally, the R^2 is 58% and individual coefficients have similar magnitude.

[Table X about here.]

The "financial" factor relates to different interest rate spreads, which is consistent with the evidence above that the liquidity factor is related to risk premia across markets.[37] The $F6$ and $F7$ factors share a similar and extremely interesting interpretation: these are "monetary conditions" factors.

[34] We also considered the Pastor-Stambaugh measure of aggregate stock market liquidity and found no relationship.

[35] Ludvigson and Ng (2009) use univariate regressions of individual series on each principal component to characterize its information content. For example, the "real" factor was labeled as such because it has high explanatory power for real quantities (e.g. Industrial Production).

[36] A significant link between liquidity and one of the principal components of LN does not necessarily require that this component predicts bond excess returns. The liquidity factor is endogenous and its loadings on the underlying macroeconomic variables is unlikely to be linear nor constant through time.

[37] LN found that the information content of the "financial" factor for excess returns is subsumed in the CP factor. Recall from Section A that the information content of the funding liquidity factor is not subsumed by the Cochrane-Piazzesi factor.

Both have highest explanatory power for the rate of change in reserves and non-borrowed reserves of depository institutions. Next, factor $F6$ has most information for the rate of change of the monetary base and the M1 measure of money stock and some information from the PCE indices. Beyond bank reserves, factor $F7$ is most informative for the spreads of commercial paper and three-month Treasury bills above the Federal Reserve funds rate. This suggests an important channel between monetary policy and the intermediation mechanism and, ultimately, with variations in the valuation of marketwide liquidity. These results are consistent with Longstaff (2004), who establishes a link between variations of RefCorp spreads and measures of flows into money market mutual funds, Longstaff et al. (2005), who document a similar link for the non-default component of corporate spreads and, finally, Chordia et al. (2005), who document that money flows and monetary surprises affect measures of bond market liquidity.

We find that the liquidity factor is also related to the "real", "inflation" factors, indicating that some of the predictability of macro factors for bond risk premium could be measured in funding markets. This may also result from the impact of the Fed's actions on funding markets. The $F5$ is a "housing activity" factor and is also significant. It contains information on housing starts and new building permits. Nonetheless, its significance appears to be limited to the early part of the sample and it is not robust to the inclusion of bid-ask spreads information (see below). Finally, the "real" and "inflation" factors are not robust to the inclusion of stock index implied volatility.

B Transaction Costs Variables

Coupon bond quotes from the CRSP data set include bid and ask prices. At each point in time, we consider the entire cross-section of bonds and compute the difference between the median and the minimum bid-ask spreads. This measures the difference in transaction costs between the most liquid bond and a typical bond. Table X presents the results from a regression of liquidity on this measure of relative transaction costs. The coefficient is positive and significant. The liquidity factor increases when the median bid-ask spread moves further away from the minimum spread. That is, on-the-run bonds become more expensive when they offer relatively lower transaction costs. The explanatory power of bid-ask information is substantial, as measured by an R^2 of 37.7%. However, there is a sharp structural break in this relationship. Most of the explanatory power and all of the statistical evidence is driven by observations preceding 1990 as made clear by Figure 1.7a.

The first break in this process coincides with the advent of the GovPX platform while the second break, around 1999, matches the introduction of the eSpeed electronic trading platform. Although transaction costs contribute to the on-the-run premium, the lack of variability since these breaks implies a lesser role in the variations of liquidity on Treasury markets.

C Aggregate Uncertainty

The valuation of liquidity should increase with higher aggregate uncertainty. We use implied volatility from options on the S&P 500 stock index as proxy for aggregate uncertainty. The S&P500 index comprises a large share of aggregate wealth and its implied volatility can be interpreted as a forward looking indicator of wealth volatility. The sample comprises monthly observations of the CBOE VOX index from January 1986 until the end of 2007. Table X presents results from a regression of liquidity on aggregate uncertainty (Regression C). The R^2 is 7.9% and the coefficients is positive but the evidence is statistically weak. Figure 1.7b shows the measures of volatility and funding liquidity until the end of 2008. Clearly, peaks in volatility are often associated with rises in liquidity valuation. The weak statistical evidence is due to the period around 2002 where very low funding liquidity value was not matched with a proportional decrease of implied volatility. In any case, the coefficient estimate suggests that a one-standard deviation shock to implied volatility raises the liquidity factor by 0.052.

D Combining Regressors

Finally, Table X reports the results from a regression combining all the economic information considered above (Regression D). The coefficient on the relative bid-ask spread decreases but remains significant. On the other hand, the information from the VIX measure is subsumed in other regressors. Its coefficient changes sign and becomes insignificant. In particular, the VIX measure is positively correlated with the stock market factor and this factor's coefficient doubles. Next, the inflation, real and housing activity factor become insignificant. However, the "monetary conditions" factors also remain significant when conditioning on transaction costs and aggregate uncertainty information.

Overall the evidence points toward two broad channels in the determination of the value of funding liquidity. First, similar to the model of Krishnamurthy and He (2008), aggregate uncertainty and aggregate wealth affect the intermediaries' ability to provide liquidity. Second, the Fed implements its monetary policy primarily through the funding market. To some extent, it also support to the stability of the financial system through that channel. Then, these policies, through their impact on funding conditions directly impact risk premium in other markets.

VII The Events Of 2008

We repeat the estimation of the model including data from 2008. Figure 1.7 presents the liquidity (Panel 1.8a) and the term structure (Panel 1.8b)factors. The latter shows a sharp increase in the cross-section of on-the-run premium. In fact, this large shock increases the volatility of the liquidity factor substantially. Looking at Figure 1.7b and 1.7a we see that this spike was associated with a large increase in the SP500 implied volatility but, interestingly, the spread between the minimum and median bid-ask spread remained stable. This supports our interpretation that the liquidity factor finds its roots in the funding market.

Adding 2008 only increases the measured impact of the common funding liquidity factor on bond risk premia. Each of the regression above leads to higher estimate for the liquidity coefficient. An interesting case, though, is the behavior of corporate bond spreads. Clearly corporate bond spreads increased sharply over that period, indicating an increase in expected returns. What is interesting is that this was the case for any ratings. Figure 1.8 compares the liquidity factor with the spread of the AAA and BBB Merrill Lynch index. In the sample excluding 2008, the estimated average impact a shock to funding liquidity was negative for AAA bonds and positive for BBB. The large and positively correlated shock in 2008 reverses this conclusion for AAA bonds. But note that AAA spreads and the liquidity factor were also positively correlated in 1998. This confirms our conjecture that the behavior of high-rating bonds is not stable and depends on the nature or the size of the shock to funding liquidity. Note that this does not affect our conclusion that corporate bond liquidity premium shares a component with other risk premium due to funding risk. Instead, it suggests that the relationship exhibits regimes through time.

VIII Conclusion

We augment the Arbitrage Free Extended Nelson-Siegel term structure model of Christensen et al. (2007) by allowing for a liquidity factor driving the on-the-run premium. Estimation of the model proceeds directly from coupon bond prices using a non-linear filter. We identify from a panel of Treasury bonds a common liquidity factor driving on-the-run premia at different maturities. Its effect increases with maturity and decreases with the age of a bond.

This liquidity factors measures the value of the lower funding and transaction costs of on-the-run bonds. It predicts a substantial share of the risk premium on off-the-run bonds. It also predicts LIBOR spreads, swap spreads and corporate bond spreads. The pattern across interest rate markets and credit ratings is consistent with accounts of flight-to-liquidity events. However, the effect is pervasive in normal times. The evidence points toward the importance of funding liquidity for the intermediation mechanism and, hence, for asset pricing. Our results are robust to changes in data set and to the inclusion of term structure information.

The liquidity factor varies with transaction costs on the secondary bond market. More importantly, we find that the value of liquidity is related to narrow measures monetary aggregates and measures of bank reserves. It also varies with measures of stock market valuations and aggregate uncertainty. The ability of intermediaries to meet the demand for immediacy depends, in part, on funding conditions and induces a large common liquidity premium in key interest rate markets. In particular, our results suggest that the behavior of the Fed is a key determinants of the liquidity premium. It remains to be seen if the impact of aggregate liquidity extends to the risk premium for stocks. In this context, the measure of funding liquidity proposed here can be used as real-time measure of liquidity premia.

IX Appendix

A Data

We use end-of-month prices of U.S. Treasury securities from the CRSP data set. We exclude callable bonds, flower bonds and other bonds with tax privileges, issues with no publicly outstanding securities, bonds and bills with less than 2 months to maturity and observations with either bid or ask prices missing. Our sample covers the period from January 1986 to December 2008. We also exclude the following suspicious quotes.

CRSP ID	Date
#19920815.107250	August 31^{st} 1987
#19950331.203870	December 30^{th}
#19980528.400000	May 30^{th} 1998
#20011130.205870	October 31^{th}
#20041031.202120	November 29^{th} 2002
#20070731.203870	May 31^{st} 2006
#20080531.204870	November 30^{th} 2007

CRSP ID #20040304.400000 has a maturity date preceding its issuance date, as dated by the U.S. Treasury. Finally, CRSP ID #20130815.204250 is never special and is excluded.

B Unscented Kalman Filter

The UKF is based on an approximation to any non-linear transformation of a probability distribution. It has been introduced in Julier et al. (1995) and Julier and Uhlmann (1996) (see Wan and der Merwe (2001) for textbook treatment) and was first imported in finance by Leippold and Wu (2003).

Given $\hat{X}_{t+1|t}$ a time-t forecast of state variable for period $t+1$, and its associated MSE $\hat{Q}_{t+1|t}$ the unscented filter selects a set of Sigma points in the distribution of $X_{t+1|t}$ such that

$$\bar{\mathbf{x}} = \sum_i w^{(i)} x^{(i)} = \hat{X}_{t+1|t}$$

$$\mathbf{Q}_x = \sum_i w^{(i)} (x^{(i)} - \bar{\mathbf{x}})(x^{(i)} - \bar{\mathbf{x}})' = \hat{Q}_{t+1|t}.$$

Julier et al. (1995) proposed the following set of Sigma points,

$$x^{(i)} = \begin{cases} \bar{\mathbf{x}} & i = 0 \\ \bar{\mathbf{x}} + \left(\sqrt{\frac{N_x}{1-w^{(0)}} \sum_x}\right)_{(i)} & i = 1, \ldots, K \\ \bar{\mathbf{x}} - \left(\sqrt{\frac{N_x}{1-w^{(0)}} \sum_x}\right)_{(i-K)} & i = K+1, \ldots, 2K \end{cases}$$

with weights

$$w^{(i)} = \begin{cases} w^{(0)} & i = 0 \\ \frac{1-w^{(0)}}{2K} & i = 1, \ldots, K \\ \frac{1-w^{(0)}}{2K} & i = K+1, \ldots, 2K \end{cases}$$

where $\left(\sqrt{\frac{N_x}{1-w^{(0)}} \sum_x}\right)_{(i)}$ is the i-th row or column of the matrix square root. Julier and Uhlmann (1996) use a Taylor expansion to evaluate the approximation's accuracy. The expansion of $y = g(x)$ around $\bar{x}$ is

$$\begin{aligned} \bar{y} &= E\left[g(\bar{x} + \Delta x)\right] \\ &= g(\bar{x}) + E\left[D_{\Delta x}(g) + \frac{D^2_{\Delta x}(g)}{2!} + \frac{D^3_{\Delta x}(g)}{3!} + \cdots\right] \end{aligned}$$

where the $D^i_{\Delta x}(g)$ operator evaluates the total differential of $g(\cdot)$ when perturbed by Δx, and evaluated at $\bar{x}$. A useful representation of this operator in our context is

$$\frac{D^i_{\Delta x}(g)}{i!} = \frac{1}{i!}\left(\sum_{j=1}^{n} \Delta x_j \frac{\partial}{\partial x_j}\right)^i g(x)\bigg|_{x=\bar{x}}$$

Different approximation strategies for $\bar{y}$ will differ by either the number of terms used in the expansion or the set of perturbations Δx. If the distribution of Δx is symmetric, all odd-ordered terms are zero. Moreover, we can re-write the second terms as a function of the covariance matrix P_{xx} of Δx,

$$\bar{y} = g(\bar{x}) + \left(\nabla^T P_{xx} \nabla\right) g(\bar{x}) + E\left[\frac{D^4_{\Delta x}(g)}{4!} + \cdots\right]$$

Linearisation leads to the approximation $\hat{\bar{y}}_{lin} = g(\bar{x})$ while the unscented approximation is exact up to the third-order term and the σ-points have the correct covariance matrix by construction. In the Gaussian case, Julier and Ulhmann (1996) show that same-variable fourth moments agree as well and that all other moments are lower than the true moments of Δx. Then, approximation errors of higher order terms are necessarily smaller for the UKF than for the EKF. Using a similar argument, but for approximation of the MSE, Julier and Uhlmann (1996) show that linearisation and the unscented transformation agree with the Taylor expansion up to the second-order term and that approximation errors in higher-order terms are smaller for the UKF.

C Arbitrage-Free Yield Adjustment Term

Christensen et al. (2007) show that the constant, $a(m)$ is given by

$$\begin{aligned}
a(m) = & -\sigma_{11}^2 \frac{m^2}{6} - (\sigma_{21}^2 + \sigma_{22}^2)\left[\frac{1}{2\lambda^2} - \frac{1-e^{-m\lambda}}{m\lambda^3} + \frac{1-e^{-2m\lambda}}{4m\lambda^3}\right] \\
& - (\sigma_{31}^2 + \sigma_{32}^2 + \sigma_{33}^2) \\
& \times \left[\frac{1}{2\lambda^2} + \frac{e^{-m\lambda}}{\lambda^2} - \frac{me^{-2m\lambda}}{4\lambda} - \frac{3e^{-2m\lambda}}{4\lambda^2} - \frac{2(1-e^{-m\lambda})}{m\lambda^3} + \frac{5(1-e^{-2m\lambda})}{8m\lambda^3}\right] \\
& - (\sigma_{11}\sigma_{21})\left[\frac{m}{2\lambda} + \frac{e^{-m\lambda}}{\lambda^2} - \frac{1-e^{-m\lambda}}{m\lambda^3}\right] \\
& - (\sigma_{11}\sigma_{31})\left[\frac{3e^{-m\lambda}}{\lambda^2} + \frac{m}{-2\lambda} + \frac{me^{-m\lambda}}{\lambda}\right] \\
& - (\sigma_{21}\sigma_{31} + \sigma_{22}\sigma_{32}) \\
& \times \left[\frac{1}{\lambda^2} + \frac{e^{-m\lambda}}{\lambda^2} - \frac{e^{-2m\lambda}}{\lambda^2} - \frac{3(1-e^{-m\lambda})}{m\lambda^3} + \frac{3(1-e^{-2m\lambda})}{4m\lambda^3}\right].
\end{aligned}$$

Table I: Summary statistics of bond characteristics

We present summary statistics of age (in months), duration (in months) and coupon (in %) for each maturity and liquidity category. *New* refers to the on-the-run security and *Old* refers to the off-the-run security (see text for details). In each case, the first column gives the sample mean and the second column gives the sample standard deviations. Coupon statistics are not reported for maturity categories of 12 months and less as T-bills do not pay coupons. End-of-month data from CRSP (1985:12-2007:12).

	Age				Duration				Coupon			
Maturity	Old		New		Old		New		Old		New	
3	12.01	9.31	1.64	0.09	3.01	0.03	4.38	0.09				
6	16.93	6.27	0.12	0.11	6.00	0.10	5.90	0.11				
9	14.45	6.05	4.42	4.88	8.89	0.11	10.00	0.40				
12	13.11	5.78	2.51	3.90	111.77	0.23	12.14	1.08				
18	28.29	11.92	6.74	0.62	17.14	0.50	16.81	0.59	7.12	2.81	6.84	3.06
24	22.90	13.45	0.33	0.52	22.56	0.59	22.68	0.72	7.11	2.82	6.74	3.13
36	24.64	10.17	4.61	6.74	32.56	1.42	32.75	2.63	7.49	2.96	7.10	2.82
48	18.42	9.57	4.42	3.00	41.95	2.30	44.17	4.40	7.38	3.03	7.25	2.86
60	29.06	21.58	2.29	3.85	50.41	3.09	51.36	3.02	7.72	2.80	7.13	2.90
84	34.41	8.61	12.51	11.82	65.85	5.04	68.71	8.45	7.74	2.56	7.55	2.63
120	14.91	18.59	4.02	7.56	84.43	8.34	85.55	9.16	7.15	2.24	7.44	2.71

Table II: Parameter estimates of transition equations.

Panel (a) presents estimation results for the AFENS model without liquidity. Panel (b) presents estimation results for the AFENS model with liquidity. For each parameter, the first standard error (in parenthesis) is computed from the QMLE covariance matrix (see Equation 1.23) while the second is computed from the outer product of scores (see Equation 1.25). End-of-month data from CRSP (1985:12-2007:12).

(a)

	$\bar{F}$	K		Σ ($\times 10^2$)	
	0.0545	0.169	0.68		
Level	(0.0136)	(0.177)	(0.42)		
	(0.0093)	(0.069)	(0.03)		
	-0.0172	0.182	0.76	0.84	
Slope	(0.0277)	(0.088)	(0.75)	(0.46)	
	(0.013)	(0.071)	(0.06)	(0.04)	
	-0.0128	0.891	-0.14	0.41	2.31
Curvature	(0.0061)	(0.860)	(1.86)	(1.64)	(0.66)
	(0.0045)	(0.283)	(0.15)	(0.17)	(0.13)

(b)

	$\bar{F}$	K		Σ ($\times 10^2$)	
	0.0576	0.198	0.85		
Level	(0.0165)	(0.165)	(0.86)		
	(0.0154)	(0.098)	(0.02)		
	-0.0167	0.222	-0.81	0.85	
Slope	(0.0092)	(0.293)	(0.85)	(0.44)	
	(0.0165)	(0.145)	(0.06)	(0.05)	
	-0.0189	0.887	0.57	0.25	2.27
Curvature	(0.0057)	(1.414)	(0.82)	(1.91)	(1.66)
	(0.0088)	(0.325)	(0.13)	(0.20)	(0.12)

	$\bar{L}$	ϕ_l	σ_l
	0.32	0.955	0.06
Liquidity	(0.42)	(0.034)	(0.066)
	(0.09)	(0.021)	(0.011)

Table III: Mean Pricing Errors and Root Mean Squared Pricing Errors

Panel (a) presents MPE and Panel (b) presents RMSPE from AFENS models with and without liquidity. The columns correspond to liquidity category where New refers to on-the-run issues and Old refers to off-the-run issues. End-of-month data from CRSP (1985:12-2007:12).

(a) Mean Pricing Errors

Mean Pricing Errors				
	Benchmark Model		Liquidity Model	
Maturity	Old	New	Old	New
3	0.015	0.026	0.003	-0.002
6	-0.002	0.02	0.018	-0.011
9	-0.031	0.026	-0.012	0.008
12	-0.041	0.037	-0.023	0.016
18	-0.064	-0.061	0.002	-0.001
24	-0.028	$4e^{-5}$	0.007	-0.005
36	0.005	0.069	0.014	-0.012
48	-0.008	0.079	-0.019	-0.003
60	0.006	0.25	0.013	0.023
84	-0.167	-0.041	0.035	-0.015
120	-0.239	0.07	-0.157	0.107
All	-0.058	0.043	-0.011	0.010

(b) Root Mean Squared Errors

Root Mean Squared Pricing Errors				
	Benchmark Model		Liquidity Model	
Maturity	Old	New	Old	New
3	0.047	0.046	0.04	0.021
6	0.035	0.041	0.038	0.023
9	0.054	0.062	0.038	0.036
12	0.072	0.078	0.057	0.049
18	0.092	0.09	0.04	0.034
24	0.063	0.085	0.056	0.062
36	0.102	0.139	0.101	0.072
48	0.171	0.183	0.161	0.088
60	0.226	0.318	0.217	0.108
84	0.327	0.298	0.235	0.176
120	0.353	0.394	0.285	0.308
All	0.177	0.197	0.145	0.089

Table IV: On-the-run Premium

Each line corresponds to a maturity category (months). The first two columns provide the average of residual differences in each category for the AFENS model with and without maturity, respectively. The last three columns display estimates of the liquidity level, $\hat{\beta}$, followed by standard errors (in parenthesis). The first standard error is computed from the QMLE covariance matrix (see Equation 1.23) while the second is computed from the outer product of scores. (see Equation 1.24). End-of-month data from CRSP (1985:12-2008:12).

Maturity	Residual Differences		$\hat{\beta}$	Standard Error	
	Benchmark	Liquidity		QMLE	MLE
3	0.0111	-0.0053	0.2642	0.0304	0.0232
6	0.0221	-0.0295	0.2837	0.0326	0.0273
9	0.0566	0.0202	0.3158	0.0370	0.0331
12	0.0783	0.0396	0.3026	0.0362	0.0335
18	0.0025	-0.0036	0.0428	0.0248	0.0352
24	0.028	-0.0117	0.2005	0.0320	0.0350
36	0.0644	-0.026	0.5325	0.0739	0.0842
48	0.0892	0.0165	0.7446	0.0945	0.0880
60	0.2477	0.0102	1.227	0.1369	0.1197
84	0.125	-0.0509	1.2174	0.1026	0.0978
120	0.3106	0.264	1	-	-

Table V: Results from off-the-run excess returns regressions

Results from predictive regression,

$$xr^{(m)}_{t+h} = \alpha^{(m)}_h + \delta^{(m)}_h L_t + \beta^{(m)T}_h F_t + \epsilon^{(m)}_{(t+h)},$$

the liquidity, L_t, and term structure factors, F_t, from the AFENS model where $xr^{(m)}_{t+h}$ is the excess returns at horizon h (months) on a bond of maturity m (years). Regressors are demeaned and divided by its standard deviation. Panel (a) contains estimates of α and Panel (b) contains estimates of δ with t-statistics based on Newey-West standard errors (h+3 lags) in parenthesis. Panel (c) presents R^2 of including or excluding [in brackets] the liquidity factor. End-of-month data from CRSP (1985:12-2007:12).

(a) Average risk premia

	Bond Maturity											
Horizon	2		3		4		5		7		10	
1	1.53	(7.07)	2.09	(11.17)	2.59	(15.00)	3.03	(18.53)	3.80	(24.86)	4.71	(33.49)
3	1.36	(4.17)	1.90	(6.64)	2.39	(8.89)	2.83	(10.89)	3.57	(14.36)	4.44	(18.90)
6	1.10	(2.67)	1.63	(4.38)	2.10	(5.89)	2.51	(7.22)	3.21	(9.53)	3.99	(12.53)
12	0.69	(1.37)	1.21	(2.59)	1.66	(3.62)	2.07	(4.50)	2.78	(6.03)	3.58	(8.07)
24	0.00	(0.00)	0.61	(0.96)	1.11	(1.67)	1.56	(2.20)	2.34	(2.94)	3.26	(3.78)

(b) Liquidity Coefficients

	Bond Maturity											
Horizon	2		3		4		5		7		10	
1	-1.39	(-2.49)	-2.27	(-2.53)	-3.01	(-2.47)	-3.61	(-2.39)	-4.52	(-2.27)	-5.42	(-2.07)
3	-1.35	(-3.28)	-2.12	(-3.14)	-2.74	(-2.97)	-3.23	(-2.84)	-3.98	(-2.64)	-4.70	(-2.34)
6	-1.25	(-4.67)	-2.00	(-4.51)	-2.59	(-4.29)	-3.07	(-4.09)	-3.84	(-3.75)	-4.69	(-3.26)
12	-0.85	(-5.47)	-1.63	(-5.63)	-2.24	(-5.63)	-2.73	(-5.57)	-3.44	(-5.18)	-4.08	(-4.15)
24	0.00	(0.00)	-0.53	(-3.24)	-0.91	(-3.23)	-1.17	(-3.27)	-1.51	(-3.29)	-1.75	(-2.91)

(c) R^2

	Bond Maturity											
Horizon	2		3		4		5		7		10	
1	4.74	[2.28]	4.65	[2.02]	4.51	[1.95]	4.34	[1.93]	4.03	[1.92]	3.55	[1.89]
3	13.56	[6.84]	13.33	[6.83]	13.07	[7.03]	12.78	[7.18]	12.07	[7.17]	10.52	[6.57]
6	24.23	[10.34]	24.50	[11.21]	24.57	[12.26]	24.61	[13.11]	24.44	[14.10]	22.92	[14.02]
12	35.36	[11.23]	37.71	[12.66]	39.24	[14.96]	40.32	[17.20]	41.46	[21.00]	40.54	[24.42]
24	0.00	[0.00]	35.53	[16.92]	31.91	[13.91]	29.46	[11.92]	26.56	[10.32]	25.82	[12.69]

Table VI: Off-the-run excess returns and funding liquidity

Results from the regressions,

$$xr^{(m)}_{t+12} = \alpha^{(m)} + \delta^{(m)} L_t + \beta^{(m)T} f_t + \epsilon^{(m)}_{(t+12)},$$

where $xr^{(m)}_{t+h}$ is the annual excess returns on a bond with maturity m (years), L_t is the liquidity factor and f_t is a vector of annual forward rates $f^{(h)}_t$ from 1 to 5 years. Regressors are demeaned and divided by their standard deviations. Panel (a) presents results using returns and forward rates directly from CRSP data but with the liquidity factor from the model. Panel (b) excludes the liquidity factor. Panel (c) excludes the liquidity factor and uses excess returns from the model. Newey-West t-statistics (in parenthesis) with 15 lags. End-of-month data from CRSP (1985:12-2007:12).

(a) Excess returns and forward rates from Fama-Bliss data with the liquidity factor

Maturity	cst	$f^{(1)}_t$	$f^{(2)}_t$	$f^{(3)}_t$	$f^{(4)}_t$	$f^{(5)}_t$	L_t	R^2
2	0.72	0.29	-1.31	1.88	0.93	-0.95	-0.78	41.65
	(3.49)	(0.49)	(-1.18)	(1.50)	(1.04)	(-1.60)	(-5.97)	
3	1.31	0.15	-2.26	4.32	0.76	-1.49	-1.55	41.66
	(3.41)	(0.14)	(-1.13)	(1.89)	(0.48)	(-1.27)	(-5.93)	
4	1.79	-0.51	-1.74	4.58	1.53	-1.85	-2.18	42.82
	(3.53)	(-0.35)	(-0.66)	(1.51)	(0.75)	(-1.13)	(-6.07)	
5	1.98	-1.51	-0.24	4.57	0.36	-0.81	-2.66	40.87
	(3.23)	(-0.84)	(-0.07)	(1.24)	(0.15)	(-0.39)	(-5.83)	

(b) Excess returns and forward rates from Fama-Bliss data

Maturity	cst	$f^{(1)}_t$	$f^{(2)}_t$	$f^{(3)}_t$	$f^{(4)}_t$	$f^{(5)}_t$	L_t	R^2
2	0.72	-0.43	-1.34	2.66	0.99	-1.53		21.04
	(2.95)	(-0.57)	(-1.06)	(1.50)	(0.95)	(-2.13)		
3	1.31	-1.27	-2.33	5.86	0.88	-2.64		19.29
	(2.87)	(-0.87)	(-1.04)	(1.77)	(0.46)	(-1.86)		
4	1.79	-2.52	-1.83	6.74	1.70	-3.46		19.86
	(2.95)	(-1.26)	(-0.62)	(1.51)	(0.67)	(-1.76)		
5	1.98	-3.96	-0.35	7.20	0.56	-2.79		18.27
	(2.71)	(-1.65)	(-0.10)	(1.35)	(0.19)	(-1.14)		

(c) Excess returns from the model and forward rates from Fama-Bliss data

Maturity	cst	$f^{(1)}_t$	$f^{(2)}_t$	$f^{(3)}_t$	$f^{(4)}_t$	$f^{(5)}_t$	L_t	R^2
2	0.66	-0.13	-1.91	2.97	0.93	-1.51		21.10
	(2.71)	(-0.17)	(-1.53)	(1.69)	(0.91)	(-2.09)		
3	1.27	-1.15	-2.04	4.97	1.19	-2.43		18.19
	(2.82)	(-0.79)	(-0.90)	(1.50)	(0.63)	(-1.73)		
4	1.74	-2.46	-1.26	6.09	1.18	-2.92		17.22
	(2.83)	(-1.24)	(-0.41)	(1.34)	(0.46)	(-1.47)		
5	2.09	-3.86	0.00	6.62	1.06	-3.12		17.15
	(2.80)	(-1.61)	(0.00)	(1.20)	(0.34)	(-1.26)		

Table VII: LIBOR rolling excess returns and funding liquidity

Results from the regressions,

$$xr_{t+h}^{(m)} = \alpha_h^{(m)} + \delta_h^{(m)} L_t + \beta_h^{(m)T} F_t + \epsilon_{(t+h)}^{(m)},$$

where $xr_{t+h}^{(m)}$ is the returns at time $t+h$ (months) on rolling investment in loans with maturity m (months), L_t is the liquidity factor and F_t is the vector of term structure factor. Each regressor is demeaned and divided by its standard deviation for interpretation. Panel (a) contains estimates of average returns. Panel (b) contains estimates of $\delta_h^{(m)}$. Newey-West t-statistics (h+3 lags) in parenthesis. Panel (c) presents R^2 from the regressions including and excluding [in brackets] the liquidity factor. End-of-month data from CRSP (1985:12-2007:12).

(a) Average Excess Returns

	Loan Maturity							
Horizon	1		3		6		12	
1	0.277	(0.347)	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)
3	0.183	(0.248)	0.265	(0.245)	0.000	(0.000)	0.000	(0.000)
6	0.062	(0.322)	0.144	(0.264)	0.239	(0.165)	0.000	(0.000)
12	-0.153	(0.615)	-0.070	(0.560)	0.029	(0.439)	0.253	(0.151)
24	-0.537	(1.120)	-0.453	(1.079)	-0.351	(0.985)	-0.120	(0.743)

(b) Liquidity Coefficients

	Loan Maturity							
Horizon	1		3		6		12	
1	0.184	(7.837)	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)
3	0.162	(7.853)	0.149	(6.364)	0.000	(0.000)	0.000	(0.000)
6	0.193	(6.139)	0.173	(6.985)	0.101	(5.699)	0.000	(0.000)
12	0.360	(5.700)	0.340	(6.364)	0.277	(7.329)	0.076	(3.695)
24	0.732	(5.578)	0.715	(5.909)	0.664	(6.395)	0.526	(7.366)

(c) R^2

	Loan Maturity							
Horizon	1		3		6		12	
1	46.4	[28.0]	0.0	[0.0]	0.0	[0.0]	0.0	[0.0]
3	44.7	[16.8]	50.6	[26.5]	0.0	[0.0]	0.0	[0.0]
6	24.7	[1.4]	30.7	[2.9]	44.8	[20.4]	0.0	[0.0]
12	29.2	[7.1]	30.3	[6.6]	32.3	[6.7]	35.2	[18.6]
24	38.8	[12.3]	38.9	[11.7]	39.4	[11.2]	41.2	[10.1]

Table VIII: LIBOR and swap spreads and funding liquidity

Results from regressions,

$$sprd_t^{(m)} = \alpha^{(m)} + \delta^{(m)} L_t + \beta^{(m)T} F_t + \epsilon_{(t)}^{(m)},$$

where $sprd_t^{(m)}$ is the spread at time t and for maturity m (months), L_t is the liquidity factor and F_t is the vector of term structure factor. Spreads are computed above the off-the-run U.S. Treasury yield curve and we use par yields to compute swap spreads. Each regressors is demeaned and divided by its standard deviation. Panel (a) presents results for LIBOR spreads. Panel (b) presents results for swap spreads. Newey-West t-statistics (3 lags) in parenthesis. Finally, R^2 from regressions including and excluding [in brackets] the liquidity factor. End-of-month data from CRSP (1985:12-2007:12).

(a) LIBOR Spreads

	1 month		3 months		6 months		12 months	
Avg Spread	0.423	(0.027)	0.422	(0.023)	0.406	(0.019)	0.429	(0.019)
$\delta_m^{(h)}$	0.183	(6.463)	0.153	(5.939)	0.106	(5.166)	0.080	(4.410)
R^2	58.4	[44.9]	59.4	[47.8]	53.2	[42.2]	53.9	[37.7]

(b) Swap Spreads

	24		60		84		120	
Avg. Spread	0.384	(0.016)	0.483	(0.018)	0.477	(0.019)	0.432	(0.020)
$\delta_m^{(h)}$	0.094	(4.556)	0.104	(4.525)	0.107	(4.395)	0.095	(3.917)
R^2	37.8	[35.4]	38.0	[34.2]	45.5	[38.6]	51.7	[38.5]

Table IX: Corporate bond excess returns and funding liquidity

Results from the regressions

$$y_t = \alpha_h^{(r)} + \delta_h^{(r)} L_t + \beta_h^{(r)T} F_t + \epsilon_{(t+h)}^{(r)},$$

where y_t is either a spread, $sprd_t^r$, observed a time t for rating r or an excess returns, $xr_{t+h}^{(r)}$ over an horizon h (months) on an investment the Corporate index with rating r, L_t is the liquidity factor and F_t is the vector of term structure factor. See Equation 1.26 for the spread panel specification. Panel (b) presents results for corporate spreads. Panel (a)) presents results for excess returns. Individual corporate bond yields are obtained from NAIC. Corporate bond returns are computed using Merrill Lynch indices obtained from Bloomberg. Spreads and excess returns are computed above the Treasury off-the-run yield curve. Each regressor is demeaned and divided by its standard deviation for interpretation. Newey-West t-statistics in parenthesis and R^2 from regressions including and excluding [in brackets] the liquidity factor. Results from Merrill Lynch indices cover the entire sample. Results from NAIC corporate bond yields is monthly from February 1996 until December 2001.

(a) Merrill Lynch Indices Excess Returns

	AAA		AA		A		BBB		HY	
Avg. Spread	3.162	(15.502)	3.130	(15.291)	3.162	(15.618)	3.204	(16.196)	3.785	(23.400)
$\delta_m^{(G)}$	-1.775	(-1.396)	-1.626	(-1.341)	-1.154	(-0.913)	0.073	(0.057)	3.117	(1.461)
R^2	4.5	[3.7]	4.9	[4.2]	4.4	[4.1]	3.4	[3.4]	6.3	[5.2]

(b) NAIC Corporate Spreads

	G1		G2		G3		G4		G5	
Avg.	1.51	(0.19)	1.65	(0.21)	2.25	(0.30)	3.38	(0.59)	3.70	(0.54)
$\delta_m^{(G)}$	-0.31	(-2.98)	-0.20	(-1.96)	-0.04	(-0.34)	0.25	(2.29)	0.26	(2.47)
R^2	3.9	[2.0]	5.7	[2.0]	6.5	[2.0]	7.0	[2.0]	7.5	[2.0]

(c) Merrill Lynch Spread Indices

	Aaa		Aa		A		Baa		HY	
Avg. Spread	0.930	(0.036)	0.976	(0.049)	1.227	(0.046)	1.856	(0.077)	5.385	(0.270)
$\delta_m^{(h)}$	0.065	(2.294)	0.060	(1.188)	0.073	(1.268)	0.119	(1.379)	0.334	(1.168)
R^2	59.5	[55.5]	31.4	[29.6]	39.6	[34.9]	49.7	[42.7]	39.2	[29.9]

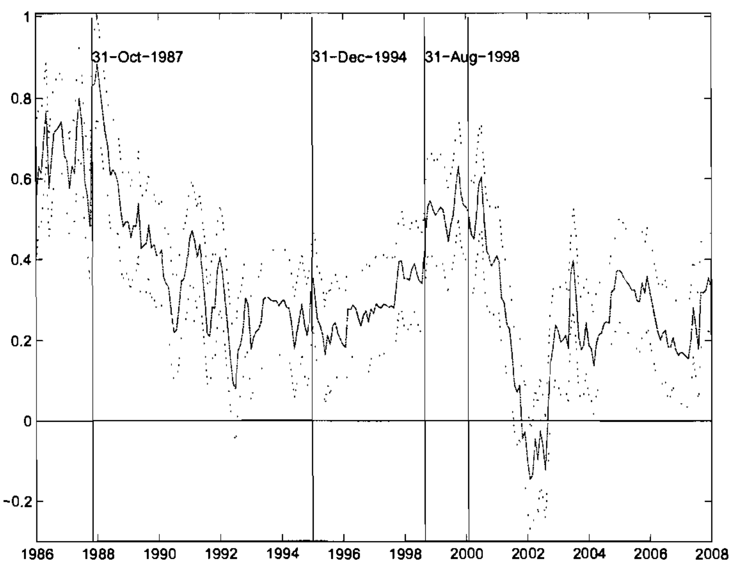
Table X: Macroeconomic determinants

Results from regressions of the liquidity factor on selected economic variables. BA is the difference between the minimum and the median bid-ask spreads across bonds on any given date. VXO is the implied volatility from S&P500 call options. $F1$ to $F8$ are principal components of macroeconomic series from Ludvigson and Ng (2009). Newey-West standard errors (3 lags) are included in parenthesis. End-of-month data (1986:01-2004:12).

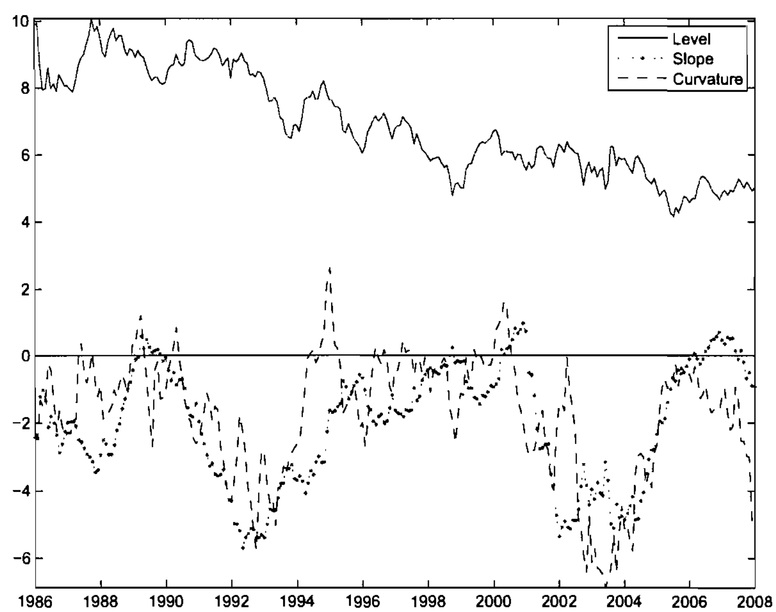
	Regressors											
Model	cst	BA	VXO	$F1$	$F2$	$F3$	$F4$	$F5$	$F6$	$F7$	$F8$	R^2
A	0.36			0.046	0.091	-0.001	0.051	0.050	-0.035	0.037	-0.030	45.0
	(16.7)			(2.13)	(5.73)	(-0.06)	(2.51)	(3.14)	(-2.34)	(2.28)	(-1.84)	
B	0.34	0.114										37.7
	(17.5)	(5.35)										
C	0.34		0.052									7.9
	(13.4)		(1.91)									
D	0.36	0.076	-0.087	0.218	0.075	0.004	0.023	0.021	-0.030	0.031	-0.059	55.5
	(19.9)	(4.13)	(0.43)	(1.40)	(4.50)	(0.37)	(1.45)	(1.35)	(-2.10)	(2.29)	(4.90)	

Figure 1.1: Liquidity and Term structure factors

Factors from the AFENS model with liquidity. Panel (a) displays the liquidity factor. The scale is in dollar. Panel (b) displays the term structure factors. The scale is in percentage. End-of-month data from CRSP (1985:12-2007:12).



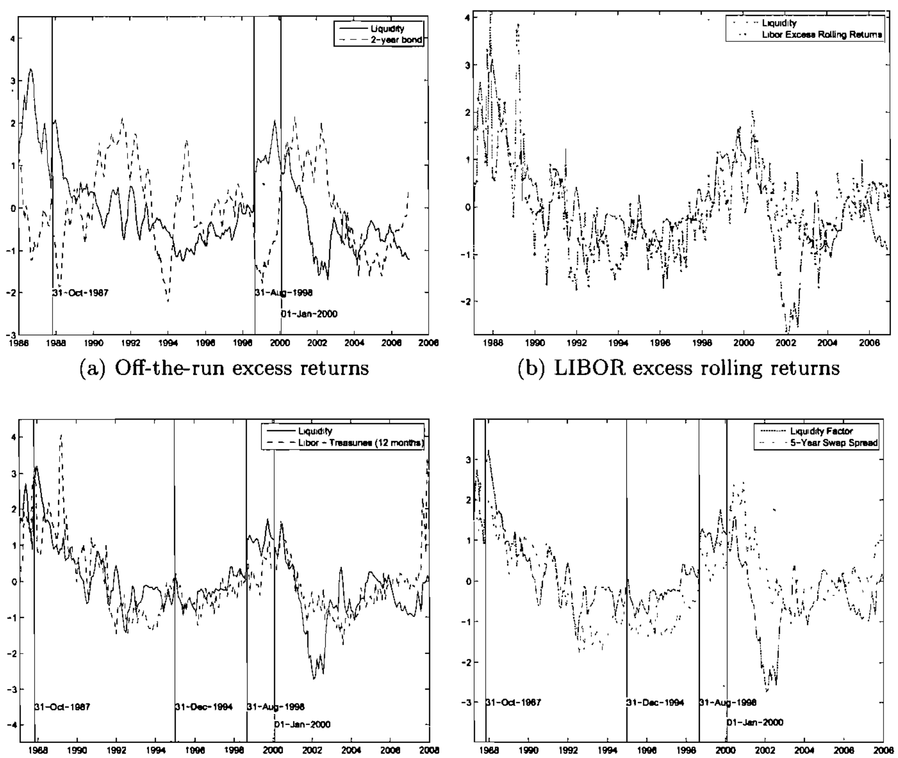
(a) Liquidity Factor



(b) Term Structure Factors

Figure 1.2: Excess returns and funding liquidity

The liquidity factor and the risk premium in different markets. Panel (a) displays annual excess returns on 2-year off-the-run U.S. Treasury bonds. Panel (b) displays annual excess rolling returns on a 12-month LIBOR loan. Panel (d) displays the spread of the 1-year LIBOR rate above the off-the-run 1-year zero yield. Panel (d) displays the spread of the 5-year swap rate. Excess returns are computed above the off-the-run Treasury risk-free rate. End-of-month data from CRSP (1985:12-2007:12).



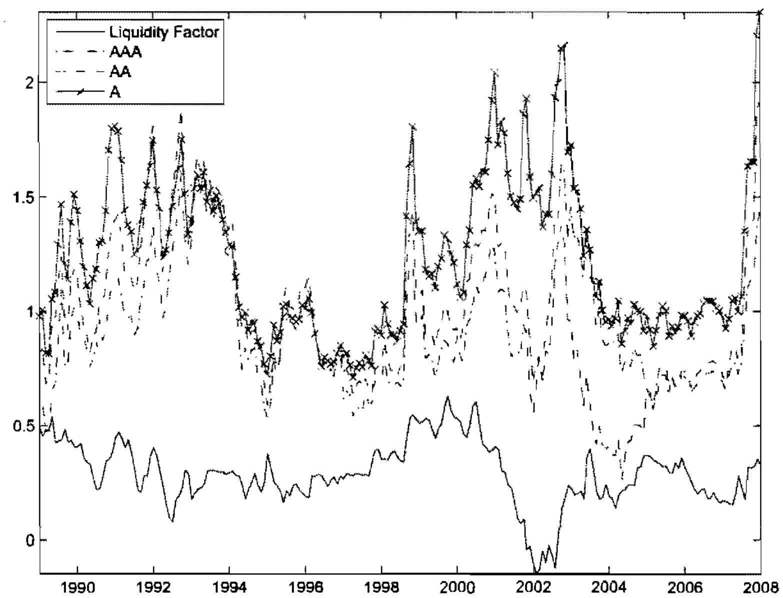
(a) Off-the-run excess returns

(b) LIBOR excess rolling returns

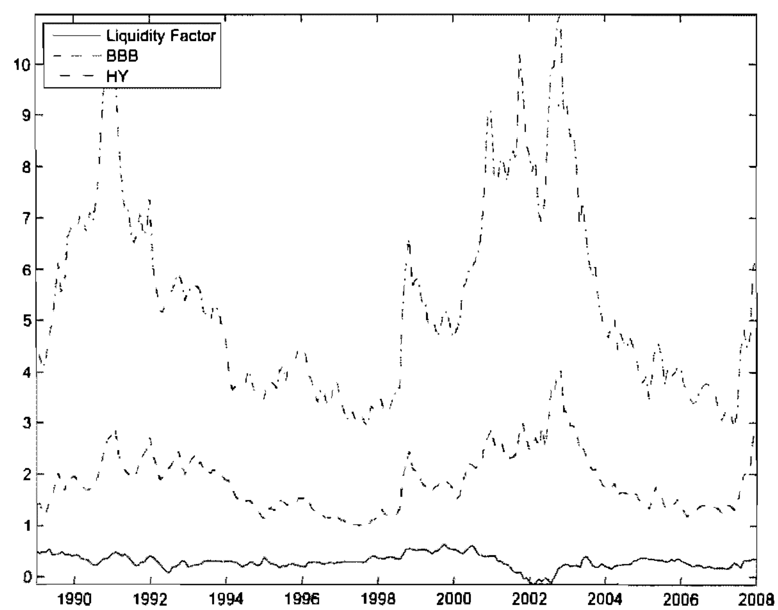
(c) 1-year LIBOR spread

(d) 5-year swap spread

Figure 1.3: Corporate spread and funding liquidity
The liquidity factor with corporate bond spreads for different ratings. Panel (a) compares with the spreads of Merrill Lynch indices for high quality bonds: AAA, AA and A ratings. Panel (b) compares with the spread of Merrill Lynch BBB and High Yield corporate bond indices. Spreads are computed above the off-the-run 10-year Treasury par yield.

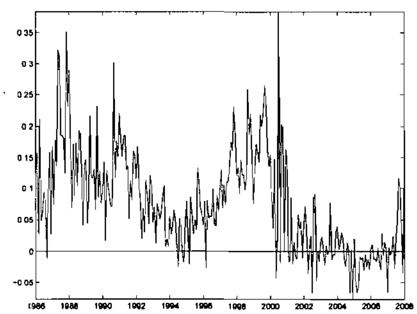


(a) Liquidity and Merrill Lynch AAA, AA, and A indices

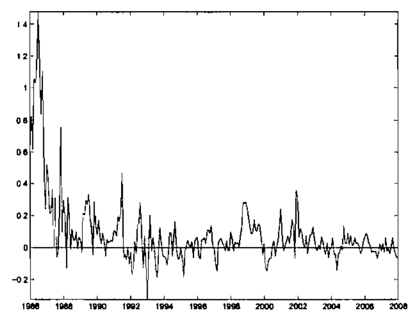


(b) Liquidity and Merrill Lynch BBB and High Yield indices

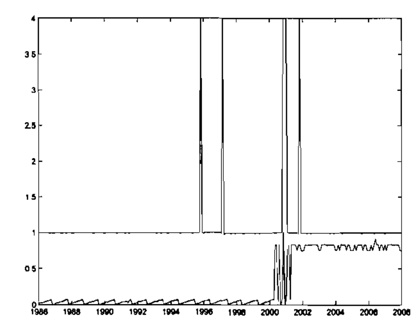
Figure 1.4: Residual Differences - Benchmark Model
Comparison of residual differences and ages for the benchmark AFENS model without liquidity. Panel (a) presents differences between the residuals (dollars) of the on-the-run and off-the-run bonds in the 12-month category. Panel (b) presents the residuals 48-month category. Panel (c) and (d) displays years from issuance for the more recent and the seasoned bonds in the 12-month and the 48-month category, respectively. End-of-month data from CRSP (1985:12-2007:12).



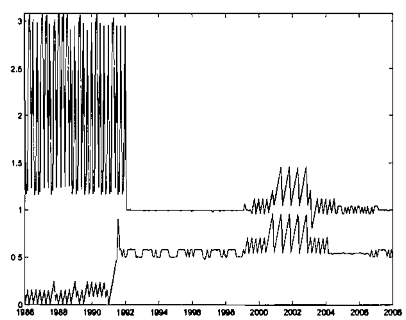
(a) 12-month residual difference



(b) 48-month residual difference

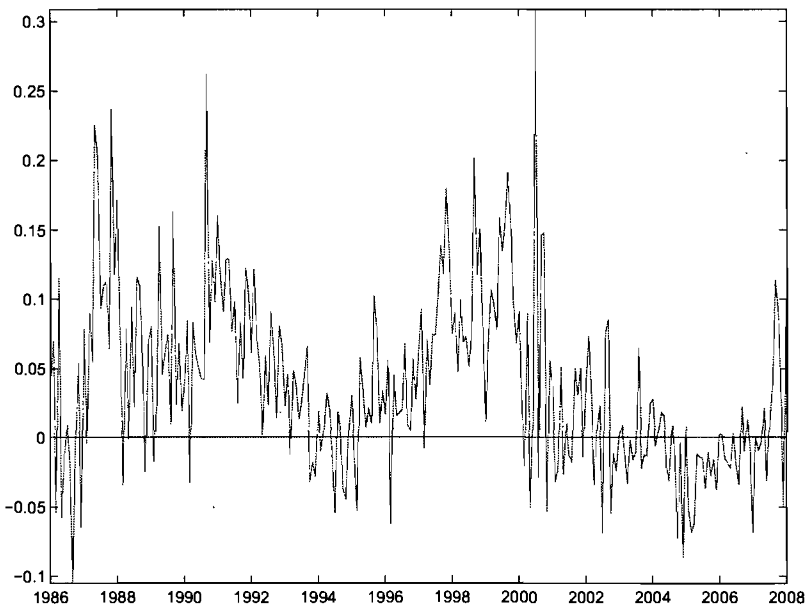


(c) 12-month ages

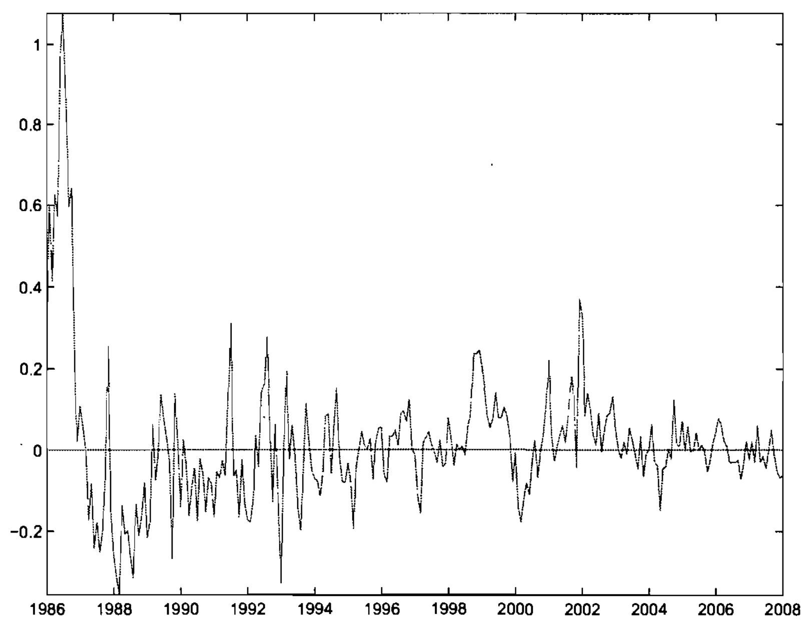


(d) 48-month ages

Figure 1.5: Residual differences - Liquidity Model
Comparison of residual differences for the AFENS model with liquidity. Panel (a) present differences between residuals (dollars) of on-the-run and off-the-run bonds in the 12-month category. Panel (b) presents differences between residuals (dollars) in the 48-month category. End-of-month data from CRSP (1985:12-2007:12).

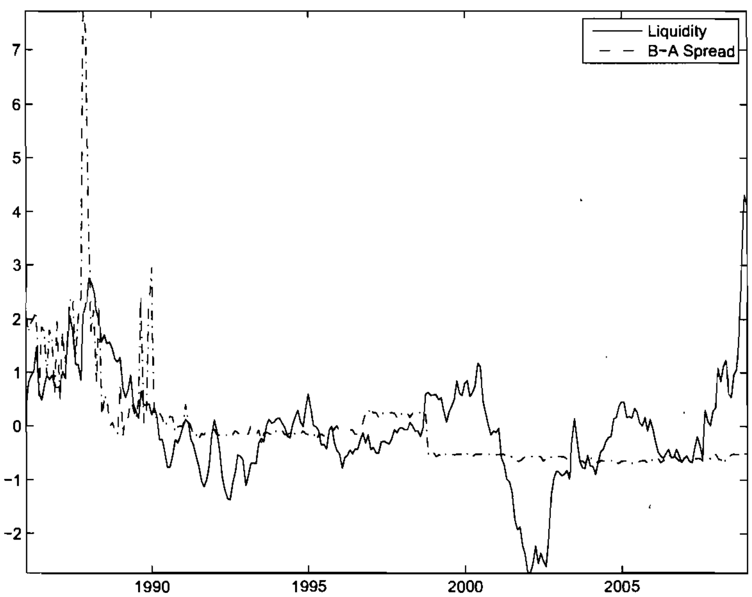


(a) 12-month residual differences

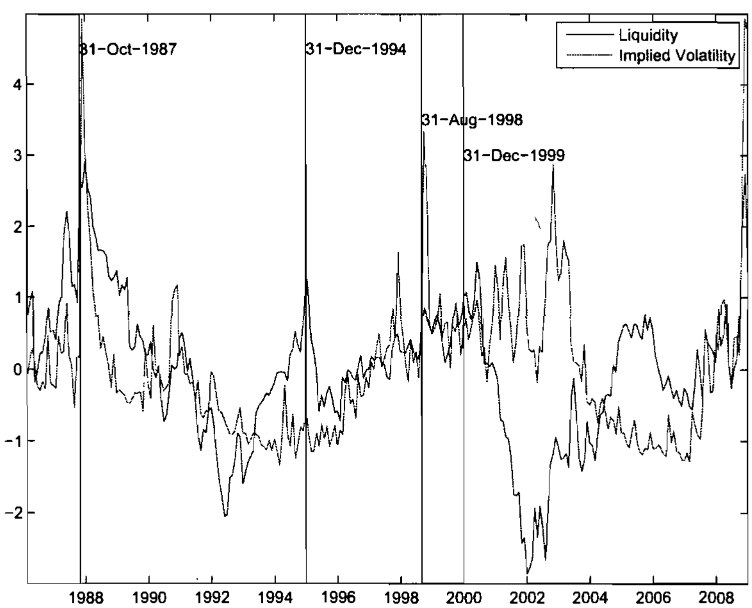


(b) 48-month residual differences

Figure 1.6: Determinants of Liquidity
Panel (a) traces the liquidity factor and the difference between the median and the minimum bid-ask spread at each observation date. Panel (b) traces the liquidity factor and implied volatility from S&P 500 call options. The liquidity factor is obtained from the AFENS model with liquidity. End-of-month data from CRSP (1985:12-2008:12)

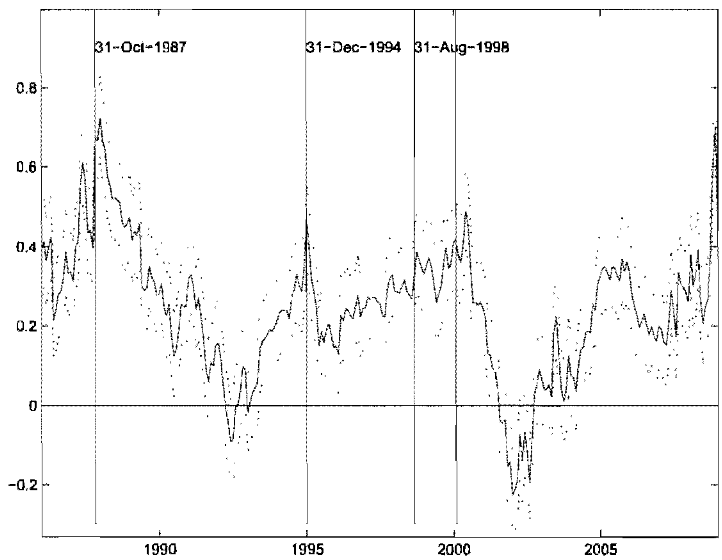


(a) Bid-Ask Spread and Liquidity

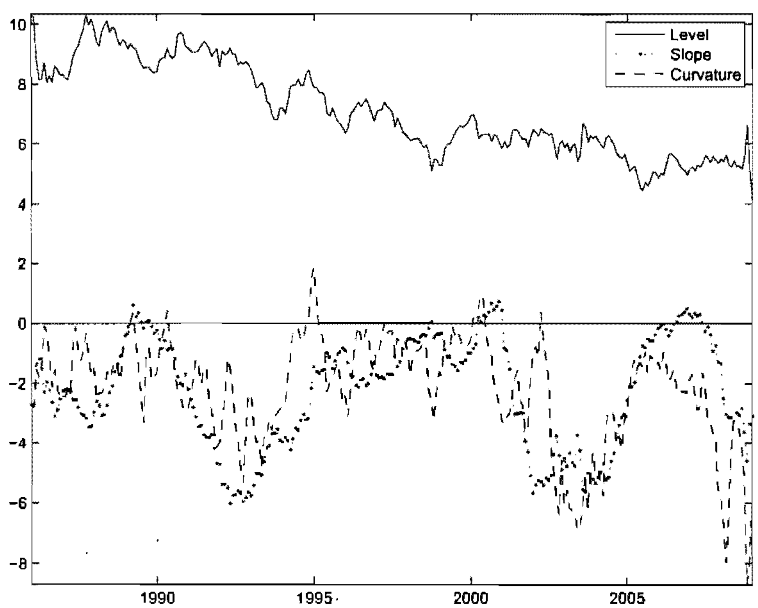


(b) Volatility and Liquidity

Figure 1.7: Liquidity and Term structure factors including 2008 Data
Factors from the AFENS model with liquidity. Panel (a) displays the liquidity factor. The scale is in dollar. Panel (b) displays the term structure factors. The scale is in percentage. End-of-month data from CRSP (1985:12-2008:12).

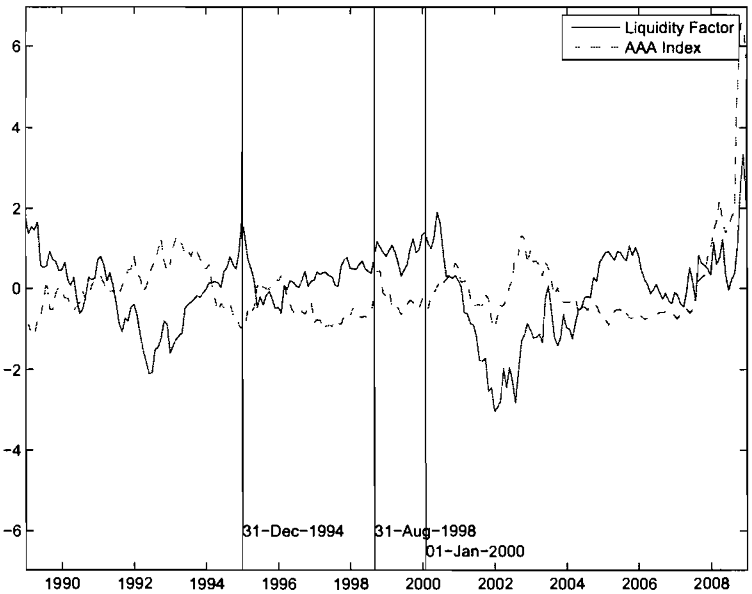


(a) Liquidity Factor

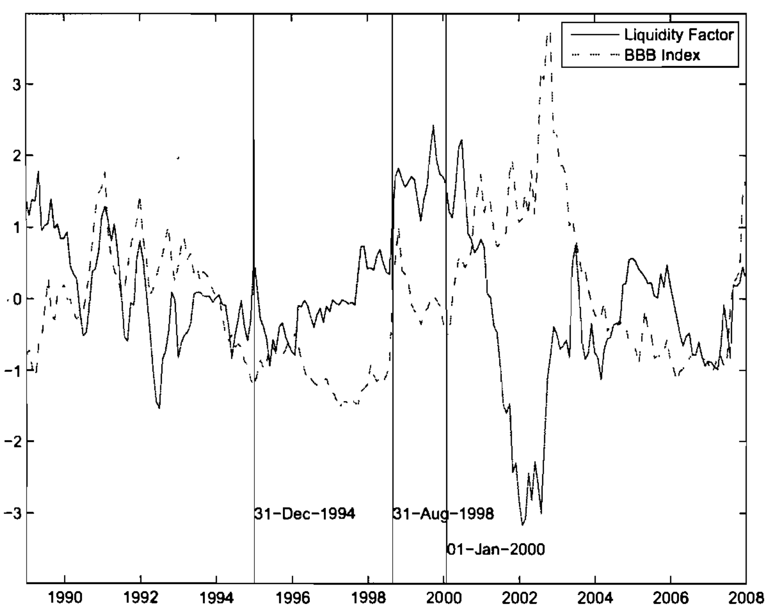


(b) Term Structure Factors

Figure 1.8: Corporate spread and funding liquidity including 2008 data
The liquidity factor with corporate bond spreads for different ratings. Panel (a) compares with the spreads of Merrill Lynch index for AAA bonds. Panel (b) compares with the spread of Merrill BBB corporate bond index. Spreads are computed above the off-the-run 10-year Treasury par yield. End-of-month data from CRSP and Merrill Lynch (1988:12-2008:12).



(a) Liquidity and Merrill Lynch AAA indiex



(b) Liquidity and Merrill Lynch BBB index

Bibliography

Acharya, V. V. and L. Pedersen (2004). Asset pricing with liquidity risk. *NBER* (W10814).

Adrian, T., E. Etula, and H. Shin (2009). Risk appetite and exchange rates. Working Paper.

Adrian, T. and H. Shin (2008). Liquidity and leverage. *Journal of Financial Intermediation*, forthcoming.

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets 5*, 31–56.

Amihud, Y. and H. Mendelson (1986). Asset pricing and the bid-ask spreads. *Journal of Financial Economics 17*, 223–249.

Amihud, Y. and H. Mendelson (1991). Liquidity, maturity and the yields on U.S. treasury securities. *The Journal of Finance 46*, 1411–1425.

Anderson, N., F. Breedon, M. Deacon, A. Derry, and G. Murphy (1996). *Estimating and interpreting the yield curve.* Series in Financial Economics and Quantitative Analysis. John Wiley & Sons.

Bakshi, G., P. Carr, and L. Wu (2005). Stochastic risk premiums, stochastic skewness in currency options, and stochastic discount factors in international economies. *Journal of Financial Economics*. forthcoming.

Bansal, R. and W. Coleman (1996). A monetary explanation of the term premium, equity premium and risk-free rate puzzle. *The Journal of Political Economy 104*, 396–409.

Bliss, R. (1997). Testing term structure estimation methods. *Advances in Futures and Options Research 9*, 197–231.

Brunnermeier, M. K. and L. Pedersen (2008). Market liquidity and funding liquidity. *Review of Financial Studies*, forthcoming.

Buraschi, A. and D. Menini (2002). Liquidity risk and specialness. *Journal of Financial Economics 64*, 243–284.

Campbell, J. and R. Shiller (1991). Yield spreads and interest rate movements: A bird's eye view. *Review of Economic Studies 58*, 495–514.

Cheria, J., E. Jacquier, and R. Jarrow (2004). A model of the convenience yields in on-the-run treasuries. *Review of Derivatives Research 7*, 79–97.

Chordia, T., A. Sarkar, and A. Subrahmanyam (2005). An empirical analysis of stock and bond market liquidity. *Review of Financial Studies 18*, 85–129.

Christensen, J., F. Diebold, and G. Rudebusch (2007). The affine arbitrage-free class of nelson-siegel term structure models. Technical Report 20, FRBSF.

Christoffersen, P., K. Jacobs, L. Karoui, and K. Mimouni (2007). Estimating term structure modesl using swap rates. McGill University.

Cochrane, J. and M. Piazzesi (2005a). Bond risk premia. *American Economic Review 95*, 138–160.

Cochrane, J. and M. Piazzesi (2005b). Reply to Dai, Singleton and Yang.

Collin-Dufresne, P., R. Goldstein, and J. Spencer Martin (2001). The determinants of credit spread changes. *The Journal of Finance 56*, 2177–2207.

Dai, Q. and K. Singleton (2000). Specification analysis of affine term structure models. *The Journal of Finance 55*, 1943–1978.

Dai, Q. and K. Singleton (2002). Expectation puzzles, time-varying risk premia and affine model of the term structure. *Journal of Financial Economics 63*, 415–441.

Dai, Q., K. Singleton, and W. Yang (2004). Predictability of bond risk premia and affine term structure models. University of Chicago.

Diebold, F. and C. Li (2006). Forecasting the term structure of government bond yields. *Journal of Econometrics 130*, 337–364.

Diebold, F., G. Rudebusch, and S. Aruoba (2006). The macroeconomy and the yield curve : A dynamic latent factor approach. *Journal of Econometrics 131*, 309–338.

Duffee, G. (2002). Term premia and interest rate forecast in affine model. *The Journal of Finance 57*, 405–443.

Duffie, D. (1996). Special repo rates. *The Journal of Finance 51*, 493–526.

Duffie, D., N. Gârleanu, and L. Pedersen (2005). Over-the-counter markets. *Econometrica 73*, 1815–1847.

Duffie, D. and R. Kan (1996). A yield-factor model of interest rates. *Mathematical Finance 6*, 379–406.

Duffie, D. and K. Singleton (1997). An econometric model of the term structure of interest-rate swap yields. *The Journal of Finance 52*, 1287–1321.

Elton, E. and T. Green (1998). Tax and liquidity effects in pricing governement bonds. *The Journal of Finance 53*, 1533–1562.

Ericsson, J. and O. Renault (2006). Liquidity and credit risk. *The Journal of Finance 61*, 2219–2250.

Fama, E. (1976). Forward rates as predictors of future spot rates. *Journal of Financial Economics 3*, 361–377.

Fama, E. (1984a). The information in the term structure. *Journal of Financial Economics 13*, 509–528.

Fama, E. (1984b). Term premium in bond returns. *Journal of Financial Economics 13*, 529–546.

Fama, E. and R. Bliss (1987). The information in long-maturity forward rates. *American Economic Review 77*, 680–692.

Fedlhütter, P. and D. Lando (2007). Decomposing swap spreads. Copenhagen Business School.

Froot, K. and P. O'Connell (2008). On the pricing of intermediated risks: Theory and application to catastrophe reinsurance. *Journal of Banking and Finance 32*, 69–85.

Gabaix, X., A. Krishnamurthy, and O. Vigneron (2009). Limits of arbitrage: Theory and evidence from the mortgage-backed securities market. *The Journal of Finance*. Forthcoming.

Gârleanu, N., L. Pedersen, and A. Pothesman (2009). Demand-based option pricing. *Review of Financial Studies*. Forthcoming.

Goldreich, D., B. Hanke, and P. Nath (2005). The price of future liquidity : Time-varying liquidity in the U.S. treasury market. *Review of Finance 9*, 1–32.

Green, R. and B. Ødegaard (1997). Are there tax effects in the relative pricing of the U.S. government bonds? *The Journal of Finance 52*, 609–633.

Grinblatt, M. (2001). An analytic solution for interest rate swap spreads. *International Review of Finance 2*, 113–149.

Gromb, D. and D. Vayanos (2002). Equilibrium and welfare in markets with financially constrained arbitrageurs. *Journal of Financial Economics 66*, 361–407.

Gurkaynak, R., B. Sack, and J. Wright (2006). The u.s. treasury curve: 1961 to present. Technical Report 2006-28.

He, Z. and A. Krishnamurthy (2007). Intermediated asset prices. Technical report.

Holmström, B. and J. Tirole (1998). Private and public supply of liquidity. *The Journal of Political Economy 106*, 1–40.

Jordan, B. and S. Jordan (1997). Special repo rates: An empirical analysis. *The Journal of Finance 52*, 2051–2072.

Julier, S. and J. Uhlmann (1996). A general method for approximating nonlinear transformations of probability distributions. Robotics Research Group, University of Oxford.

Julier, S., J. Uhlmann, and H. Durrant-Whyte (1995). A new approach for filtering nonlinear systems. *The Proceedings of the American Control Conference 3*, 1628–1632. Seattle.

Kiyotaki, N. and R. Wright (1989). On money as a medium of exchange. *The Journal of Political Economy 97*, 927–954.

Krishnamurthy, A. (2002). The bond/old-bond spread. *Journal of Financial Economics 66*, 463–506.

Krishnamurthy, A. and Z. He (2008). Intermediary asset prices. W14517.

Krishnamurthy, A. and J. Vissing-Jorgensen (2007). The demand for treasury debt. W12881.

Kyle, A. and W. Xiong (2001). Contagion as a wealth effect. *The Journal of Finance 56*, 1401–1440.

Lagos, R. (2006). Asset prices and liquidity in an exchange economy. Technical Report 373, Federal Reserve Bank of Minneapolis.

Leippold, M. and L. Wu (2003). Design and estimation of multi-currency quadratic models. *European Finance Review 7*, 47–73.

Liu, J., F. Longstaff, and R. Mandell (2006). The market price of risk in interest rate swaps: The roles of default and liquidity risks. *Journal of Business 79*, 2337–2359.

Longstaff, F. (2000). The term structure of very short-term rates : New evidence for the expectation hypothesis. *Journal of Financial Economics 58*, 397–415.

Longstaff, F. (2004). The flight-to-liquidity premium in U.S. treasury bond prices. *The Journal of Business 77*, 511–526.

Longstaff, F., S. Mithal, and E. Neis (2005). Corporate yield spreads: Default risk or liquidity? new evidence from the credit-default swap market. *The Journal of Finance 60*, 2213–2253.

Ludvigson, S. and S. Ng (2009). Macro factors in bond risk premia. *Review of Financial Studies*. forthcoming.

Luttmer, E. (1996). Asset pricing in economies with frictions. *Econometrica 64*, 1439–1467.

Nashikkar, A. and M. Subrahmanyam (2006). Latent liquidity and corporate bond yield spreads. Stern School of Business.

Nelson, C. and A. Siegel (1987). Parsimonious modeling of yield curves. *The Journal of Business 60*, 473–489.

Pastor, L. and R. Stambaugh (2003). Liquidity risk and expected stock returns. *The Journal of Political Economy 111*, 642–685.

Piazzesi, M. (2005). *Affine term structure models.* in Handbook of Financial Econometrics. Elsevier.

Shi, S. (2005). Nominal bonds and interest rates. *International Economic Review 46*, 579–612.

Shi, S. (2006). Viewpoint: a microfoundation of monetary economics. *Canadian Journal of Economics 39*, 643–688.

Shiller, R. (1979). The volatility of long-term interest rates and expectation models of the term structure. *The Journal of Political Economy 87*, 1190–1219.

Shleifer, A. and R. Vishny (1997). The limits of arbitrage. *The Journal of Finance 52*, 35–55.

Startz, R. (1982). Do forecast errors of term premia really make a difference between long and short rates. *Journal of Financial Economics 10*, 323–329.

Svensson, L. (1985). Money and asset prices in a cash-in-advance economy. *The Journal of Political Economy 93*, 919–944.

Tauchen, G. and H. Zhou (2006). Realized jumps on financial markets and predicting credit spreads. Technical Report 2006-35, Federal Reserve Board.

Vayanos, D. (2004). Flight to quality, flight to liquidity, and the pricing of risk. *NBER* (W10327).

Vayanos, D. and P.-O. Weill (2006). A search-based theory of the on-the-run phenomenon. *The Journal of Finance.* forthcoming.

Wan, E. and R. V. der Merwe (2001). *Kalman filtering and neural network*, Chapter The unscented Kalman filter. Wiley and Sons.

Warga, A. (1992). Bond returns, liquidity and missing data. *The Journal of Financial and Quantitative Analysis 27*, 605–617.

Chapitre 2

Term Structure Of Monetary Expectations

Term Structure of Monetary Policy Expectations

Jean-Sébastien Fontaine
Bank of Canada

Abstract

Prices of Federal funds futures contracts are widely used to measure anticipations of monetary policy in the U.S. but a time-varying risk premium blurs their information content. This paper proposes a joint no-arbitrage model of the Federal Reserve policy function, of Federal funds rate futures and of LIBOR rates. The policy function follows a pure step process and has a Taylor rule interpretation. It is driven by the current policy rate, a macro factor, which captures the state of the economy as perceived by the monetary authority, and a liquidity factor, which captures deviations between the risk premium implicit in LIBOR and futures rates, respectively. The model extracts the expectation component common to both markets and offers real-time measures of policy anticipations. Empirically, combining LIBOR and futures rates leads to substantial improvements of policy rate forecasts at horizons up to one year. In practice, the second factor is the most important predictor of risk premium variations and, hence, an important contributor to forecasts of future Fed funds rate. I interpret deviations between futures and LIBOR rates as the reflection of forward-looking hedging activities.

Discussions with René Garcia, and its support, were instrumental for the progress of this paper. I gratefully acknowledge financial support from the Institut de Finance Mathématique de Montréal and the Banque Laurentienne. *Address for correspondence:* Bank of Canada, 234 Wellington, Ottawa, Canada.
e-mail: [redacted]

I Introduction

Prices of Federal funds futures are widely used to measure anticipations of future monetary policy in the US. Indeed, futures-based forecasts are unbiased predictors of future Target rates at short horizons (Hamilton (2009)). However, the presence of a time-varying and counter-cyclical risk premium blurs market anticipations at horizons beyond 2 months (Rudebusch (2006), Piazzesi and Swanson (2006)). This paper proposes a joint no-arbitrage model of the Federal Reserve [the Fed] monetary policy response function, of Federal funds [Fed funds] futures rates and of the term structure of London Interbank Offered Rates [LIBOR]. The Target for the overnight Fed funds rate follows a step process with jumps occurring upon FOMC meetings and where state variables drive the conditional distribution of jumps. This is a discrete-time analog to Piazzesi (2005) but extended to futures markets and allowing for a different liquidity premium in LIBOR and futures markets.

The main contribution of this paper is twofold. First, it combines interest rates with information from futures prices to achieve a sharper decomposition of rates into policy expectations and a risk premium. LIBOR and futures rates provide non-overlapping information on the future path of the Target rate[1], even in a risk-neutral world. Moreover, as discussed below, their risk premia differ. This provides further identification of the common expectation component in LIBOR and futures rates. Empirically, I estimate the model at the daily frequency and find that model-implied forecasts significantly outperform the common regression-based approach. Moreover, model-implied forecasts can be extended beyond the the maturity of the longest available futures contract.

Second, the model measures the risk premium specific to each market. Fed funds futures require no exchange of principal and have daily margin requirements that mitigate credit exposures. In contrast, LIBOR loans carry substantial default risk since they are fully funded and uncollateralized. Needless to say, participation in the LIBOR market is limited. Furthermore, investments in LIBOR loans cannot be reversed as easily, especially in periods of turmoil. Then, LIBOR rates include a compensation for liquidity and credit risk. I introduce a liquidity factor[2] which captures variations in

[1]The Fed target the overnight Fed funds rate. The effective overnight rate is a weighted average of rates on brokered unsecured loans between large banks. The LIBOR rates provide the natural term structure associated with the overnight market as they corresponds to the rates at which large bank are prepared to lend to each other on the London EuroDollar market.

[2]Hereafter, I refer to this factor as a liquidity factor and the risk premia specific to

the relative compensation for liquidity and credit risk between these markets. Empirically, this factor is the most important predictor of futures excess returns and, ultimately, leads to significantly improved policy rate forecasts. I interpret these small but informative deviations between futures and LIBOR rates as liquidity pressures in the futures market revealing the information of market participants hedging their interest rate exposures (e.g. banks). This supports similar results obtained by Piazzesi and Swanson (2006) who link excess returns to variations of commercial banks' open interest.

The paper is also a contribution to the literature on the measurement of monetary shocks (see Christiano et al. (1998)). The recursive identification scheme of structural VAR is hard to justify when including financial data: can we tell whether interest rates react to policy or whether the reverse hold? A more sensible approach is to consider monetary policy actions and interest rates as jointly determined. This is precisely what an integrated model of policy and the term structure delivers. The model provides closed-form densities of policy changes at any horizon and the no-arbitrage restrictions ensure that forecasts at different horizons are consistent with each other. As noted by Hamilton and Jordà (2002), changes in monetary policy anticipations contain a size and a timing component. A decision by the FOMC that does not accord with expectations built in market prices can be interpreted as a surprise in timing, a surprise in size, or both.

Related literature

This paper is closely related to Piazzesi (2005) but differs in some key dimensions. Using information from the futures market provides better identification of the expectation and risk premium components of interest rates. Moreover, combining assets and allowing for different risk compensations uncover an important forward-looking component. Finally, the asymmetric policy function introduces variations in the higher-order moments of the Target rate and, hence, further variations in the risk premium. Hamilton and Jordà (2002) also exploit the discrete nature of Target rate changes but do not impose no-arbitrage restrictions. They model the probability of a change as a conditional hazard model while the size of a change follows an ordered response model. Hamilton and Jordà conclude that current interest rate spreads are essential to model the hazard rate at short horizons. However, predictions of the policy rate at longer horizons are difficult because

LIBOR loans as a liquidity premium. This is in part for simplicity but also because the results below suggests that in this sample, variations in the spread are mainly due to the price impact of hedging activities. See below for further discussion.

spreads are themselves difficult to predict. In contrast, a joint model of the policy rule and of the term structure can use current price information to forecast consistently at all horizons.

Krueger and Kuttner (1996) were first to consider predictions of future monetary policy based on futures contracts. Since, Gurkaynak et al. (2007) have shown that Fed funds futures and eurodollar futures outperform other market instruments when forecasting future U.S. policy rates. Hamilton (2009) finds very little predictability of price changes for near-term futures contracts. He concludes that daily changes in futures prices accurately reflect changes in market's expectations of future monetary policy. In contrast, the evidence presented here suggests that part of these changes may be attributable to the price impact of hedging demand. This is consistent with Piazzesi and Swanson (2006). They documented that variations of banks' open interests in the futures market are closely related to excess returns on futures contracts. Moreover, Piazzesi and Swanson show that macroeconomic indicators, such as a recession dummy and changes in non-farm payroll employment are important predictors of futures excess returns. Hence, futures prices include a significant, time-varying and pro-cyclical risk premium (see also Sack (2004)). This implies that the bias in Target rate forecasts tends to increase as we enter a recession, precisely at a time when the central bank is likely to cut interest rates. While measures of employment and economic activity are available only at low frequencies, the forecasts presented here can be updated daily to provide risk-adjusted measures of future policy anticipations.

Kuttner (2001) uses futures to measure unanticipated Target changes and document the link with interest rate changes at different maturities. This contrasts with the usual result that total Target changes have a weak impact on the term structure. Similarly, Bernanke and Kuttner (2005) use futures to document the impact of unanticipated policy actions on stock prices. Recently, Rudebusch (2006) uses futures at longer horizons to argue that monetary policy in the U.S. does not follow a partial adjustment rule but, instead, adjusts the Target rapidly following the arrival of new information. In all cases, the conclusion relies on the implicit assumption that variations of the risk premium are unimportant. Alternatively, Cochrane and Piazzesi (2002) measure policy shocks from unanticipated changes in the Target Fed funds rate.

The Federal funds market is covered by a large literature.[3] Hamilton

[3]See Hamilton (1996) and Stigum (2004) for detailed discussion of the Fed funds mar-

(1996) introduces a mixed-normal ARCH model of the Fed funds rate that highlights the importance of jumps, of time-varying volatility (see also Das (2002)).[4] Jumps are more likely on Wednesday and on Thursday and on days of FOMC meetings. The former transitory day-of-the-week effects are related to microstructure of the Fed funds market while the latter FOMC meeting jumps are persistent. Johannes (2004) provides conclusive evidence that these persistent jumps affect the evolution of other short-term rates.

The rest of the paper is organized as follows. Section II introduces the model and the state variables. I also derive the associated (affine) conditional multivariate Laplace transform. Next, Section III introduces the stochastic discount factor and shows that the Laplace transform remains within the same family under the risk-neutral measure. This leads to closed-form solutions for risk-free interest rates, LIBOR rates and futures rates. Section IV summarizes the data and the Quasi-Maximum Likelihood estimator while Section V presents the estimation results. In particular, I use the unscented Kalman filter to handle non-linearities in futures prices. Section VI evaluates the forecasting implications of the model, relative to the usual regression-based approach. Also, I measure the information content of the liquidity and of the macroeconomic factor for excess returns on futures. Section VII concludes.

II Model

A Short Rate Dynamics

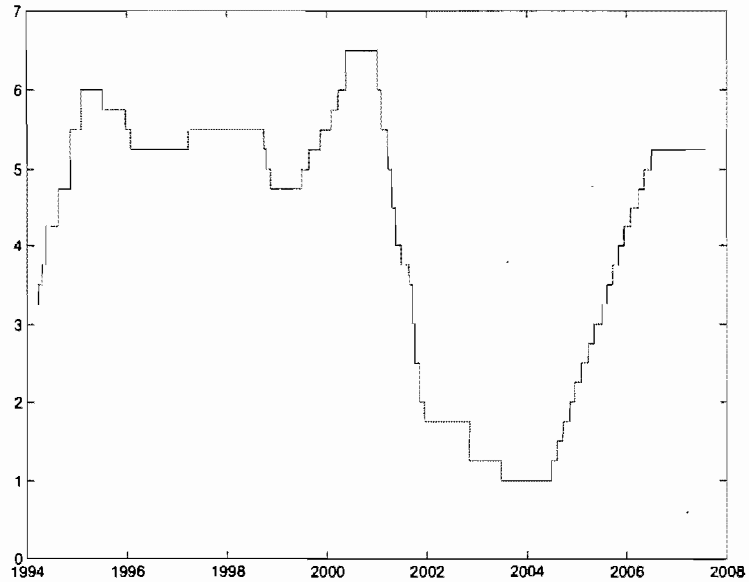
The overnight Fed funds rate is the main policy tool used by the Federal Reserve to reach its long-term objectives. Figure 2.1 draw the Target rate. Since 1994, the FOMC, which determines the Target overnight rate, has announced each change publicly and has initiated almost all of its policy changes following a scheduled meeting. Due to this operating procedure, the path of the Target Fed funds rate, r_t, traces a step function through time.[5] This suggests the following representation of the policy rule

ket.

[4] Also, the Fed funds rate tends to decrease over the 2-week reserve maintenance period before a surge on settlement day (i.e. it is not a martingale over the reserve maintenance period) and this surge is accompanied by a sharp rise in volatility on the settlement day. This pattern of volatility over the maintenance period has been confirmed in more recent data by Barolini et al. (2002).

[5] This has not been always the case. The uncertainty associated with the Fed's policy caused continuous variations of the overnight rate as new economic information arrived.

Figure 2.1: Target Federal Funds Rate (1994-2007)



$$
\begin{aligned}
r_{t+1} &= r_0 + \Delta \sum_{i=1}^{t+1} \left(n_i^u - n_i^d \right) \\
&= r_t + \Delta \left(n_{t+1}^u - n_{t+1}^d \right),
\end{aligned} \tag{2.1}
$$

for $t \geq 0$ and where $\Delta = 0.25\%$ is the fixed increment. Conditional on time-t information, n_{t+1}^u and n_{t+1}^d are independent Poisson random variables with jump intensities

$$
\begin{aligned}
\lambda_{t+1}^u &= \lambda + \lambda_u^T \left(X_{t+1}^* - \bar{X} \right) \\
\lambda_{t+1}^d &= \lambda - \lambda_d^T \left(X_{t+1}^* - \bar{X} \right),
\end{aligned} \tag{2.2}
$$

respectively, on scheduled FOMC meeting days. A small number of policy actions occurred following unscheduled meetings. I allow for the arrival of these rare events but with constant intensities λ_0^u and λ_0^d.

B State Variables

When deciding on the course of the Target rate, the committee knows the previous Target, r_{t-1}, and its spread with the effective[6] overnight rate. The committee also observes a wealth of macroeconomic and financial infor-

[6]The effective (i.e. market) Fed funds rate is published by the Board of Governors of the Federal Reserve System (see statistical release H.15). The effective rate is an average of brokered transactions weighted by the amount of overnight funds traded.

mation captured by a first latent variable, z_t. Conditional on the occurrence of a FOMC meeting, the distribution of a Target rate change is driven by the distance of each state variable relative to its long-term mean.

In the following, estimation is based on observations of LIBOR and Fed funds futures rates. In the absence of arbitrage opportunities, the Fed funds futures rates is related to the path of futures Fed funds rates as

$$F(t,n) = E_t\left[M_{t,t+T_n} D_n^{-1} \sum_{i=T_n-D_n}^{T_n} (r_{t+i} + s_{t+i})\right]$$

where $M_{t,t+T_n}$ is the Stochastic Discount Factor [SDF[7]] , T_n is the number of days until the end of the reference month and D_n is the number of days in that month. Similarly, the price of risk-free zero coupon bond with maturity m is given by

$$D^{rf}(t,m) \equiv E_t\left[M_{t,t+m}\right].$$

Finally, I introduce a liquidity factor that further discounts LIBOR loans. The price of a LIBOR loan with maturity m is

$$D^{l}(t,m) \equiv E_t\left[M_{t,t+m} \exp\left(-\sum_{i=0}^{m-1} -l_{t+i}\right)\right],$$

where l_t captures the spread between Fed funds futures and LIBOR rates unexplained by anticipations of futures policy rates and the associated risk premium. Then, the information set of the FOMC previous to any policy decision is summarized by the vector X_{t+1}^{*} of state variables,

$$X_{t+1}^{*} \equiv \left[r_t\ s_{t+1}\ z_{t+1}\ l_{t+1}\right]^T, \tag{2.3}$$

where the time-subscript on the Target rate differs. For private agents, following a policy announcement, if any, the vector of state variables, X_t, is

$$X_t \equiv \left[r_t\ s_t\ z_t\ l_t\right]^T. \tag{2.4}$$

The spread between Target and effective rates, s_t, is included in the state variable for the purpose of computing futures prices. However, in the following I assume that $\lambda_{u,s} = \lambda_{d,s} = 0$ and, thus, that the Fed does not consider the current spread in its evaluation of the appropriate policy stance. The next section describes the dynamics of the remaining state variables. Their

[7]The SDF and the associated risk-neutral probability measure are discussed below.

interpretation will be discussed below.

C State Dynamics

The state variables driving the policy function in Equation (2.1) and (2.2) follow standard autoregressive processes. First, the spread, s_t, of the effective overnight rate above the Target rate follows a process with jumps that captures its leptokurtic distribution,

$$s_{t+1} = \mu_s + \phi_s s_t + \sigma_s \epsilon^s_{t+1} + J^s_{t+1} n^s_{t+1}, \tag{2.5}$$

where $\epsilon^s_t \sim N(0,1)$. The jump term follows a compound Poisson distribution with number of jumps $n^s_{t+1} \sim P(\lambda_s)$ and jump size $J^s_{t+1} \sim N(\nu_s, \omega^2_s)$, conditionally on the number of jumps. These random variables are mutually *i.i.d.*. Next, the latent factors, z_t and l_t, follow

$$z_{t+1} = \mu_z + \phi_z z_t + \sigma_z \epsilon^z_t \tag{2.6}$$

$$l_{t+1} = \mu_l + \phi_l l_t + \sigma_l \epsilon^l_t \tag{2.7}$$

where ϵ^z_t and ϵ^l_t are mutually *i.i.d.* standard gaussian shocks.

D Conditional Laplace Transform

Conditionally on X^*_t, Target rate changes have a Skellam distribution. This corresponds to the distribution of the difference of two Poisson processes for which the density and characteristic functions are known explicitly (Skellam (1946)). Then, the conditional distribution of the state vector, X_t, can be characterized through its conditional multivariate Laplace transform,

$$\begin{aligned} T(u, X_{t+1}) &\equiv E_t\left[\exp\left(u^T X_{t+1}\right)\right] \\ &= \exp\left(A(I_{t+1}, u) + B(I_{t+1}, u)^T X_t\right) \end{aligned} \tag{2.8}$$

where u belongs to $\mathbb{R}^4$ and I_t is an indicator function equal to 1 if an FOMC meeting is scheduled at $t+1$ and 0 otherwise.[8] Coefficients $A(\cdot)$ and $B(\cdot)$ are given in Appendix A.

E Discussion

Modeling the Target rate as a two-sided step process is motivated by the data. However, information about the underlying economic state variables arrives following smooth autoregressive processes. This information is summarized by the conditional intensities of "up" and "down" jumps that

[8]This model belong to the class of discrete-time compound autoregressive (CAR) processes introduced by Darolles et al. (2002).

drive the distribution of future Target changes. This specification allows for time-variation of all conditional moments of the Target rate. In fact, the relationship between the jump intensities and conditional moments of Target changes provides interesting interpretations. The first three conditional moments of Target changes are

$$\begin{aligned} E_t[r_{t+1} - r_t] = &\quad \Delta(\lambda_u + \lambda_d)^T (X^*_{t+1} - \bar{X}) \\ Var_t[r_{t+1} - r_t] = &\quad \Delta^2 (2\lambda + (\lambda_u - \lambda_d)^T (X^*_{t+1} - \bar{X})) \\ skew_t[r_{t+1} - r_t] = &\quad (\lambda_u + \lambda_d)^T (X^*_{t+1} - \bar{X}) / ((2\lambda + (\lambda_u - \lambda_d)^T (X^*_{t+1} - \bar{X})))^{3/2}. \end{aligned}$$

First, in the stationary state where each state variable is equal to its unconditional mean, the distribution of Target rate changes is symmetric around zero with variance $2\lambda\Delta^2$. Next, consider the conditional expectation of a policy change. It has an interpretation similar to a Taylor rule. The expected change depends on the deviations of each state variable from its long-run mean. Suppose that high values of the macro factor, z_t, are associated with higher than average inflation or employment. Then, given the Fed's mandate, we expect interest rates to rise when the macro factor, z_t, is above its long-run average, and, conversely, to decrease when it is below its long-run average. That is, $\lambda_u + \lambda_d$ should be positive. But note that parameters of this Taylor rule, λ_u and λ_d, cannot be identified separately using the conditional mean equation only. However, the variance and skewness of the policy rate depend on the difference $\lambda_u - \lambda_d$. These higher-order moments are time-varying whenever the policy rule is asymmetric and $\lambda_u \neq \lambda_d$. In the case where policy is more aggressive in recessions (i.e. $\lambda^d_{u,z} > \lambda^u_{d,z}$) the variance of policy changes decreases when conditions deteriorate and its distribution is skewed toward the left. Then, estimating a linear symmetric policy rule when the true policy rule is asymmetric is likely to produce cyclical forecast errors. But conditional moments are difficult to estimate in practice and the asymmetry of the policy rule is hard to measure. Extending the model to include asset prices provides identification of these parameters.

Before we proceed to complete the model note that some restrictions are necessary to ensure the stationarity of the Fed funds rate. Intuitively, the intensity of an "up" change in the Target rate must decrease and the intensity of a "down" change must increase when the Target rate increases relative to its long-term average, $\bar{r}_t$. In other words, if $\lambda_{u,r}$ and $\lambda_{d,r}$ are negative then the policy function induces reversion of the Target rate toward its long-term average, $\bar{r}$. I also assume that s_t, z_t and l_t are stationary and that ϕ_s, ϕ_z

and ϕ_l belong to the unit interval $(-1, 1)$.

III Asset Pricing

This section provides a pricing kernel and completes the specification of the asset pricing model. I derive the conditional dynamics under the risk-neutral measure and compute explicit prices of futures and LIBOR rates.

A Stochastic Discount Factor

In the standard endowment economy, obtaining a pricing kernel from first principles requires the specification of preferences over consumption paths and a stochastic process for endowment. I choose instead to use the following exponential-affine specification of the stochastic discount factor [SDF], M_{t+1},

$$M_{t+1} = \exp\left(-r_t - s_t - A(I_{t+1}, \delta_t) - B(I_{t+1}, \delta_t)^T X_t + \delta_t^T X_{t+1}\right) \qquad (2.9)$$

where δ_t is a vector of prices of risk.

Allowing for time-varying prices of risk is motivated by empirical evidence. Piazzesi and Swanson (2006) document the counter-cyclical variations in the risk premium of Fed funds futures.[9] Note that the model allows for two sources of risk premium variations. First, the risk premium may vary with the prices of risk. Second, the risk premium may vary with the volatility and higher-order moments of Target changes. This is so even if the price of risk is constant. Intuitively, the SDF is of the same form as the SDF of economies with power utilities. In this case the risk premium depends on the variance, skewness and excess kurtosis of future Target rates (see e.g. Polimenis (2006)). Then, to the extent that interest rate risk is priced, variations in the higher moments of the policy rate lead to variations of the interest rate risk premium.

Following Duffee (2002), the price of risk vector is affine in the state vector

$$\delta_t = \delta_0 + \delta_1 X_t, \qquad (2.10)$$

where δ_0 is a $K \times 1$ vector and δ_1 is $K \times K$ matrix. This representation is consistent with absence of arbitrage since we have $E_t\left[M_{t+1}\right] = \exp\left(-r_t - s_t\right)$, the price today for a dollar received tomorrow in the Federal funds market. This paper focuses on variations of the risk premium at the business cycle frequency, which is captured by variations in z_t. Then, I only allow the price

[9] But variations of the risk premium appear minimal for the shortest maturities. See Hamilton (2009).

of risk of z_t to vary through time. That is,

$$\delta_1 = [0\ 0\ \delta_{1,z}\ 0]^T, \tag{2.11}$$

with

$$\delta_{1,z}^T = [\delta_{1,zr}\ \delta_{1,zs}\ \delta_{1,zz}\ \delta_{1,zl}]. \tag{2.12}$$

B The Price Generating Function

Asset prices can be derived from the price of the payoff $\exp(u^T X_{t+m})$ for some maturity m. This is a generating system of the set of all payoffs (Duffie et al. (2000) and Gourieroux et al. (2002)). In particular, the safe payoff is obtained when $u = 0$. Assuming that no arbitrage opportunity exists, the discounted value of the generating payoff is the price generating function,

$$\begin{aligned}\Gamma(u,t,m) &= E_t\left[M_{t,t+m}\exp\left(u^T X_{t+m}\right)\right]\\ &= E_t\left[M_{t+1}\Gamma(u,t+1,m-1)\right],\end{aligned}$$

where I define $M_{t,t+i} \equiv M_{t+1}\cdots M_{t+i}$.[10] Consider first the case $m = 1$ which delivers the conditional Laplace Transform under the risk-neutral probability

$$\begin{aligned}\Gamma(u,t,1) = T^Q(u,X_{t+1}) &= E_t\left[M_{t,t+1}\exp\left(u^T X_{t+1}\right)\right]\\ &= \exp\left(A^Q(I_{t+1},u) + B^Q(I_{t+1},u)^T X_t\right),\end{aligned}$$

with coefficients given in Appendix B. The conditional distribution of the state vector is the same under the risk-neutral measure but with risk-adjusted parameters. In particular, the Target rate is Skellam under both measures. Extending to longer maturity $m > 1$, the price generating function has the following exponential-affine solution

$$\Gamma(u,t,m) = \exp\left(c_0(u,t,m) + c(u,t,m)^T X_t\right), \tag{2.13}$$

where coefficients satisfy (see appendix B),

$$\begin{aligned}c_0(u,t,m) &= c_0(u,t+1,m-1) + A^Q(I_{t+1}, c(u,t+1,m-1)))\\ c(u,t,m) &= B^Q(I_{t+1}, c(u,t+1,m-1)),\end{aligned} \tag{2.14}$$

with initial conditions $c_0(u,t,0) = 0$ and $c(u,t,0) = u$ given by $\Gamma(u,t,0) =$

[10] Also, $M_{t,t+1} = M_{t+1}$ and by convention $M_{t,t} = M_t$=1.

$\exp(u^T X_t)$.

C Computing Recursions

The recursions in Equation 2.14 have a time and a maturity dimension because the FOMC meeting schedule changes over time. That is, all future meetings get closer by one day every day. At first, it would seem that a different recursion must be computed to match each (t, m) pair, implying a dramatic increase in computing costs. Fortunately, there is a way around this. The key is to note that future meeting dates are known in advance and, therefore, that coefficients $A(\cdot)$ and $B(\cdot)$ are not stochastic. Then, for a given date, and for given parameter values, we can compute coefficients for the price of an asset maturing on this date but for past observation dates[11] since the recursions are increasing in m but decreasing in t. As an example, consider the price, at some date $t+h$, of an asset maturing on that day. Its price is $\exp(u^T X_{t+h})$, and the coefficients, $c_0(u, t+h, 0)$ and $c(u, t+h, 0)$ correspond to the initial values in the recursions. Next, consider the price of that asset on the previous day. Its maturity is now one and the coefficients, $c_0(u, t+h-1, 1)$ and $c(u, t+h-1, 1)$, are given directly from the recursions above. We can then work our way back until we reach t, the first date in the sample where an asset matures at time-$t+h$. Finally, varying the maturity date, t, provides us with all the needed coefficients[12].

D Risk-Free Discount Bonds

The price at time-t of a risk-free discount bond with maturity m is easily obtained by setting $u = 0$, that is

$$\begin{aligned} D^{rf}(t, m) &\equiv E_t\left[M_{t,t+m}\right] \\ &= \Gamma(0, t, m) \\ &= \exp\left(d_0^{rf}(t, m) + d^{rf}(t, m)^T X_t\right) \end{aligned} \tag{2.15}$$

where $d_0^{rf}(t, m) \equiv c_0(0, t, m)$ and $d^{rf}(t, h) \equiv c(0, t, m)$ for $m \geq 0$.

[11] Another approach that reduces computing cost is to assume a constant time interval between meetings beyond the nearest schedule meeting date, as in Piazzesi (2005). However, the use of Fed funds futures contracts make this approximation problematic as it may place some future meetings in the wrong month, implying severe mispricing of the corresponding futures contracts.

[12] This implies that some recursions must be started for some date $t+h$ beyond the end of the sample. Coefficients are discarded as we proceed backward in time until we reach the last observation date of the sample.

E LIBOR Loans

LIBOR loans are short-term unsecured inter-bank loans. Like risk-free rates, LIBOR rates reflect anticipations of future Target rates and a compensation to lenders for the uncertainty surrounding future rates. However, as discussed above, I allow for a further term that drives a difference between the term premium implicit in LIBOR rates and futures rates. The price, $D^l(t,m)$, of a LIBOR loan with maturity m is,

$$\begin{aligned} D^l(t,m) &\equiv E_t\left[M_{t,t+m}\exp\left(-\sum_{i=0}^{m-1} l_i\right)\right] \\ &= \exp\left(d_0^L(t,m) + d^L(t,m)^T X_t\right), \qquad (2.16) \end{aligned}$$

with coefficients given in the Appendix. The interpretation of l_t is straightforward. It represents the extra yield, relative to the risk-free asset, required by investors to hold LIBOR loans. This extra component may be due to barriers to entry, capital constraints of liquidity providers and arbitrageur, or counterparty default risk. Ultimately, these frictions due to market power, illiquidity or default risk prohibit arbitrageurs from exploiting (risk-adjusted) differences across market prices. This reduced-form approach borrows from Grinblatt (1995) and Duffie and Singleton (1997) but my focus is on the identification of policy anticipations and I do not try to distinguish between the liquidity and the credit components of LIBOR spreads.

F Federal Funds Futures

A futures contract delivers, at the end of a given reference month, the difference between the contracted futures rate and the average overnight Fed funds rate realized during that reference month, $\bar{r}_t$. With no loss of generality, I standardize the notional of the contract to 1. The time-t price, $P(t,n)$, of a Fed funds futures contract for the reference month n is

$$P(t,n) = 100 - F(t,n) \times 3600,$$

where the Futures rate[13] $F(t,n)$, is the discounted expectation of the monthly average of overnight rates,

$$\begin{aligned} F(t,n) &= E_t\left[M_{t,t+T_n^*} D_n^{-1} \sum_{i=T_n-D_n}^{T_n} r_{t+i}\right] \\ &= D_n^{-1} \sum_{i=T_n-D_n}^{T_n} E_t\left[M_{t,t+T^*} r_{t+i}\right] \\ &= D_n^{-1} \sum_{i=T_n-D_n}^{T_n} f(t,i,T_n^*), \end{aligned} \tag{2.17}$$

where T_n is the number of days until the end of the reference month and D_n is the number of days in that month. In practice, due to weekends or holidays, the settlement date, $t+T_n^*$, may not coincide with the last day of the month, $t+T_n$. In the following, I simplify the presentation and set $T^* = T$.

Equation (2.17) shows that we need only consider a single day in the reference month. I define the rate, at time-t, of a singleton futures contract, $f(t,h,T)$, for the reference day $t+h$ and settling at date $t+T$ as

$$f(t,h,T) \equiv E_t\left[M_{t,t+T} r_{t+h}\right], \tag{2.18}$$

$$= \frac{\partial}{\partial u} E_t\left[M_{t,t+T} \exp(u r_{t+h})\right]\Big|_{u=0}, \quad 0 \leq h \leq T, \tag{2.19}$$

which can be computed explicitly from the price generating function (see Appendix B). The singleton futures rate is

$$\begin{aligned} f(t,h,T) = \exp&\left(d_0^{rf}(t+h, T-h) + c_0(u^*, t, h)) + c(u^*, t, h)^T X_t\right) \\ &\times \left[c_0'(u^*, t, h) C_r + X_t^T c'(u^*, t, h) C_r\right], \end{aligned} \tag{2.20}$$

where $u^* = d(t+h, T-h)$ and $C_r = [1\ 1\ 0\ 0]^T$. The coefficients $c_0'(\cdot)$ and $c'(\cdot)$ satisfy recursions that can be obtained as the derivatives with respect to u of Equation (2.14). That is,

$$\begin{aligned} c_0'(u,t,h) &= c_0'(u,t+1,h-1) + A'^Q(I_{t+1}, c(u,t+1,h-1))c'(u,t+1,h-1) \\ c'(u,t,h) &= B'^Q(I_{t+1}, c(u,t+1,h-1))c'(u,t+1,h-1), \end{aligned} \tag{2.21}$$

where the derivatives of the Laplace transform coefficients are taken with

[13] In practice we observe P(t,n). The Futures rate is computed as F(t,n) = (100 - P(t,n))/3600. I use the 30/360 convention throughout.

respect to the second argument and given in the appendix. The initial conditions can be obtained by noting that $f(t,0,T) = D^{rf}(t,T)r_t$ which implies that $c_0'(u^*,t,0) = 0$ and $c'(u^*,t,0) = I$ for any t.

IV Data and Method

A Data

I construct a sample of daily Target and effective overnight Fed funds rates from January 1st, 1994 to the end of July 2007. The Federal Reserve Bank of New York publishes overnight Fed funds rates on its web site. I include LIBOR rates at monthly maturities of 1 to 12 months and Fed funds futures rates for monthly horizons of 1 to 6 months available from Datastream. Futures contracts at horizons beyond 6 months exhibit low liquidity for most of the sample and are excluded. In this sample, Target rate changes were announced publicly by the FOMC following each meeting and each Target change was a multiple of 25 basis points [bps]. Finally, I exclude the period following July 2007. Since then behavior of LIBOR and futures rates are qualitatively different this change is left for further research.

B Summary Statistics

The properties of LIBOR and futures rates show that these assets offer different compensations for risk. Table B presents summary statistics for LIBOR rates, LIBOR forward rates and futures rates across maturities. Panel (a) shows that the average term structure of LIBOR rates is upward sloping, from 4.32% to 4.70%, revealing a positive average term premium. The term structure of LIBOR volatilities is almost flat, ranging from 1.76% to 1.78%, but with slight hump-shape at maturity of 5 months. Next, Panel (b) presents statistics of LIBOR *forward* rates. These forward rates are more comparable to futures rates since they cover the same future time period. The average term structure of forward rates is steeper, from 4.32% to 5.07% and the hump shape in volatilities is more pronounced. Panel (c) clearly shows that futures rates are on average, lower, less volatile and exhibit a flatter term structure than forward LIBOR rates.

Figure 2.2 compares the time series of forward LIBOR rates and futures rates at maturities of 2, 4 and 6 months. Again, forward rates are higher than futures rates. However, while variations in the levels of forward and futures rates seem to agree about the future behavior of the Target rate, their difference exhibits variations through time. Table B and Figure 2.2 show that futures and LIBOR rates carry different risk premia. This highlights the potential of using both markets to identify their common expectation component.

C State-Space Representation

The joint conditional likelihood is available in closed form but two of the state variables are unobserved. I formulate the estimation problem as a state-space system and use a Kalman-based algorithm to evaluate the likelihood function and obtain filtering estimates of the latent factors. The state transition equation is given by Equations 2.1, 2.5, 2.6 and 2.7. The measurement vector is obtained by stacking LIBOR rates and futures rates. The (non-linear) measurement can then be written as

$$Y_t = \Upsilon(t, X_t) + \epsilon_t, \tag{2.22}$$

where pricing errors, $\epsilon_{i,t}$, are i.i.d. mean zero gaussian random variables with standard deviation ω_i. The non-linearity comes from the pricing equation for futures. The overall measurement equation stacks the yield vector, Y_t with the Target rate and the effective spread. The extended measurement vector is then $\tilde{Y}_t = [r_t \; s_t \; Y_t]^T$.

Because futures rates are non-linear, the Kalman filter is not applicable. Moreover, the optimal filter is not available in closed form. I use the Unscented Kalman Filter [UKF], which is based on a recent method for calculating statistics of a random variable which undergoes a nonlinear transformation (see Appendix E). This approach has a Monte Carlo flavor but the sample is drawn according to a deterministic algorithm. It reduces the computational burden considerably, relative to simulation-based methods, but provides greater accuracy than linearization. The UKF has been used recently in the term structure literature by Leippold and Wu (2003) and Fontaine and Garcia (2008). Also, Christoffersen et al. (2007) show that the UKF improves filtering and estimation performance in the context of a term structure model of swap rates.

Table I: Means (μ) and standard deviations (σ) of LIBOR rates and forward LIBOR rates for maturities from 1 to 12 months as well as futures rates for horizons of 1 to 6 months. Data from 01:01:1994 to 31:07:1996.

(a) LIBOR Rates

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
μ_{Lib}	4.32	4.37	4.41	4.44	4.47	4.50	4.54	4.57	4.60	4.63	4.67	4.70
σ_{Lib}	1.76	1.77	1.78	1.78	1.78	1.78	1.78	1.78	1.78	1.77	1.77	1.77

(b) Forward Rates

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
μ_{For}	4.32	4.41	4.49	4.53	4.61	4.66	4.73	4.79	4.84	4.95	5.01	5.07
σ_{For}	1.76	1.78	1.80	1.80	1.80	1.80	1.79	1.79	1.79	1.78	1.77	1.74

(c) Futures Rates

	1	**2**	**3**	**4**	**5**	**6**
μ_{Fut}	4.18	4.21	4.25	4.28	4.33	4.37
σ_{Fut}	1.74	1.75	1.75	1.76	1.76	1.75

D Quasi-Maximum Likelihood Estimation

Given the filtering procedure, a Quasi-Maximum Likelihood [QML] estimator is feasible. The joint log-likelihood is given by

$$\begin{aligned} L(\theta;Y) &= \sum_{t=1}^{T} \log\left(f\left(\tilde{Y}_t|\tilde{Y}_{t-1};\theta\right)\right) \\ &= \sum_{t=1}^{T} \log\left(f(r_{t+1}-r_t|\hat{X}^{*}_{t+1|t})f(s_{t+1}|s_t)f(Y_t|\hat{X}_{t+1|t})\right) \end{aligned}$$

where all model parameters are grouped in the vector θ.[14] The following constraints were imposed on the parameter space. First, I impose $\lambda_{r,u} \leq 0$ and $\lambda_{r,d} \leq 0$ to ensure stationarity of the Target rate process. Similarly, ϕ_s, ϕ_z and ϕ_l must remain within the unit circle. Also, the jump component of the effective rate is well-defined only if $\lambda_{J,s} \geq 0$. I also impose that $\lambda_{z,d} \geq 0$ and $\lambda_{z,u} \geq 0$ because their signs cannot be identified separately from the sign of z_t.

More importantly, λ_t^u and λ_t^d must remain non-negative so that the distribution of Target jumps remains well defined. These constraints cannot be easily imposed on the parameter space as they can only be checked recursively as we filter the state variables. In practice I impose that $\hat{\lambda}_t^i = \max\left(0, \lambda_t^i\right)$. This leaves the state variables unrestricted but constrains the policy function. The restriction is reasonable. As λ_t^i approaches zero, the probability distribution of the corresponding jump n_t^i approaches the trivial distribution with a unit mass at zero. When it reaches zero, the policy function becomes one-sided and can then be summarized as a Poisson distribution.

The constrained QML estimator is given by

$$\begin{aligned} \hat{\theta}^{QML} &= \underset{\theta}{\operatorname{argmax}}\ L(\theta;Y) \\ s.t. &\quad \theta \in S \end{aligned}$$

where $S \subseteq \mathbb{R}^K$ and we have $\hat{\theta} \sim N(\theta_0, T^{-1}\Omega)$ for some true parameter value, θ_0, in the interior of the parameter space. Finally, as in Piazzesi (2005), λ_0^u and λ_0^d are poorly estimated because of the lack of policy changes outside of scheduled FOMC meetings. I calibrate them so that the distribution of

[14] I set $f(x_1|x_0) = f(x_1)$ and use the unconditional mean and covariance to initiate the Kalman recursion. Estimation is carried out using the active-set algorithm from the IMSL Fortran optimization library.

policy moves is symmetric outside of scheduled FOMC meetings, that the probability of a 50 basis points move matches its empirical probability (i.e. 3 occurences) and, finally, that the variance of Target changes matches the sample variance. Results are robust to the choice of calibration strategy.

V Estimation Results

This section reports estimation results. The key conclusions are the following. First, the model provides a good fit of LIBOR and futures rates. Second, parameter estimates indicate that the policy rule is not symmetric. In particular, the policy function exhibits mean-reversion but the Fed's mean-reverting behavior is stronger when the interest rate is below its long-run average than when it is above. Also, filtering results show that the macroeconomic factor leads business cycle fluctuations in Target rates and aggregates forward-looking information about monetary policy. Finally, the liquidity factor increases in periods where financial markets were tense, reaching peaks in the summer of 1998, and around the turn of the millennium. Still, the liquidity and macro factors exhibit rich co-movement patterns. Their sample paths exhibit numerous cases where the two factors display sharp changes but of opposite signs. In practice, this allows futures rates to lead LIBOR rates and suggests that the spread between futures and LIBOR rates contains forward-looking information.

Parameter values are reported in Table II. Panel (a) displays estimated values for the parameters driving the Target rate change intensities. Panel (b) display price of risk parameters. Panel (c) displays estimated values for parameters driving the dynamics of the spread, the macroeconomic factor and of the liquidity factor. Both latent factors are very persistent, but this is not surprising given the daily observation frequency and the persistence of interest rates. Finally, Panel (d) displays estimated standard deviations of pricing error at each maturity for LIBOR and futures rates.

A Policy Function

The estimated policy function is asymmetric. In our sample, which cover much of the Greespan's era, the Fed exhibited a stronger tendency to raise interest rates when the Target rate was relatively low then to decrease it when it was relatively high. Similarly, it showed a stronger tendency to raise rates as economic conditions, as measured by z_t, improved, then to decrease rates when conditions deteriorated. Moreover, the observed asymmetry induces substantial variations in the (conditional) variance of Target changes. In particular, variance is high and skewness is negative when economic conditions are below their long-term averages, and the reverse is true when economic conditions are relatively good. Together, the estimates suggest that the Fed acted more rapidly, and that the uncertainty surrounding future policy was lower, when it faced improving economic conditions. This has important implications for the interpretation of existing, symmetric, Taylor-rule estimates. In particular, forecasts based on linear symmetric rules will exhibit cyclical performance variations.[15]

B Risk Premium

The prices of policy rate risk is positive. That low interest rates are associated with bad states of the world is likely to be caused by the endogenous response of the Fed to economic conditions. The price of liquidity risk is also positive. Then, the difference between the risk compensations implicit in futures and LIBOR rates co-vary positively with economic conditions. This is a first indication that, in this sample, variations in l_t are linked to the relative illiquidity of the LIBOR market. Indeed, it seems unlikely that its price of risk would be positive if l_t was a measure of the default risk for the average LIBOR market participant.[16]

Also, the average price of macro risk is positive. Not surprisingly, lower values of z_t are also associated to bad states of the world. While positive on average, the price of macro risk varies through time. It is lower when the policy rate is relatively high and, presumably, the Fed has room to lower interest rates in the advent of a worsening economic outlook. Also, all else constant, the price of macro risk decreases when the liquidity factor is high, and that futures rates are relatively low compared to LIBOR rates. This is consistent with an increase of l_t being a signal of improving economic conditions. Together with a positive price of risk for the liquidity factor, this suggests that these transitory deviations between LIBOR and futures rates, generally thought as unimportant, in fact incorporate important forward-looking information. This will be confirmed in the forecasting exercises below.

[15]The asymmetry in the conditional mean may be due to the asymmetric response of the economy to monetary policy. See Garcia and Schaller (1999). Also, the conditional variance of future policy changes may be due to an asymmetric pattern of economic uncertainty. In any case, forecasts based on linear symmetric homoscedastic rule will be affected.

[16]As mentioned before, this does not preclude that variations in default risk also affect LIBOR rates. What the results suggest is that variations in the wedge between LIBOR and futures risk premium were primarily caused, ultimately, by their relative illiquidity *in this sample*. Clearly, the credit component is an important factor in the recent credit crunch.

Table II: QML parameter estimates

Parameter estimates for the joint Target, LIBOR and futures model from a daily sample covering January 1994 to July 2007. Panel (a) displays parameter estimates of the policy function. Panel (b) displays price of risk parameters. Panel (c) displays parameters of the state variable dynamics and Panel (d) displays measurement error standard deviation parameters. In each case, standard errors based on the Hessian of the likelihood function evaluated at the maximum is provided in parenthesis. The estimate for λ is 0.1669 and its standard error is (0.034).

(a) Parameter of Target Rate Intensities

	r	s	z	l
λ_d	-1.067×10^4	-	8.759×10^{-2}	1.339×10^5
	(1.91×10^2)	-	(1.27×10^{-3})	(1.47×10^4)
λ_u	-2.101×10^3	-	4.837×10^{-2}	2.753×10^5
	(8.56×10^2)	-	(3.86×10^{-3})	(5.88×10^3)

(b) Price of risk parameters

	r	s	z	l
δ	6.694×10^4	-4.679×10^2	3.491×10^{-1}	8.354×10^3
	(1.99×10^3)	(7.12×10^1)	(1.27×10^{-3})	(7.43×10^2)
$\delta_{1,z}$	-2.187×10^{-2}	0	1.437×10^{-2}	-6.459×10^4
	(6.64×10^{-4})	0	(1.99×10^{-4})	(2.74×10^2)

(c) State Dynamics Parameters

	μ	ϕ	σ	λ	ν	ω
s_t	3.176×10^{-7}	0.2851	4.409×10^{-6}	3.223×10^{-8}	7.400×10^{-6}	0.260
z_t	0	0.9978	0.3765	-	-	-
	0	(2.82×10^{-4})	(1.79×10^{03})	-	-	-
l_t	3.556×10^{-8}	0.9921	2.133×10^{-7}	-	-	-
	(1.49×10^{-10})	(3.19×10^{-4})	(3.54×10^{-9})	-	-	-

(d) Standard Deviations of Measurement Errors ($\times 10^{-6}$)

Months	1	2	3	4	5	6	7	8	9	10	11	12
LIBOR	3.01	2.01	1.58	1.09	0.71	0.37	0.25	0.25	0.31	0.25	0.15	0.29
	(0.0037)	(0.0025)	(0.0020)	(0.0014)	(0.0093)	(0.0057)	(0.0043)	(0.0038)	(0.0041)	(0.0035)	(0.0040)	(0.0048)
Futures	1.22	1.46	1.84	2.30	2.61	2.69	-	-	-	-	-	
	(0.0015)	(0.0018)	(0.0023)	(0.0028)	(0.0032)	(0.0033)	-	-	-	-	-	

C LIBOR rate loadings

Figure 2.3 displays the loadings of LIBOR rates on each factor for maturities from 1 day to 1 year. Not surprisingly, parameter estimates imply that the current policy rate is a level factor. A rising Target rate lifts all LIBOR rates but the impact is slightly less for the longest rates due to mean-reversion. Next, both the liquidity and the macroeconomic factors carry positive loadings. That is, LIBOR rates rise whenever the state of economy improves or the liquidity factor rises. The macroeconomic factor (Panel (c)) is a slope factor with no impact at the shortest maturity (except on FOMC meeting dates) but quickly rising with maturity. In contrast, the liquidity factor's loading is one, by construction, at a maturity of one day. It reaches a maximum at around 3.5 at a maturity of 180 days and slowly decreases afterwards. (Panel (d)). Finally, the spread between the Target and the Effective rate only affects yields for maturities up to a few days (see Panel (b)). This reflects the spread's highly transitory nature and should not come as a surprise given the effort of the Fed to meet any expected deviations from the Target with open market operations.

D State Variables

Figure 2.4 displays the filtered time series of state variables. Panel (a) shows the path of the Target rate between 1994 and July 2007. Panel (b) shows deviations of the Effective rate from Target. More interesting is the path of the macroeconomic factor (Panel (c)), which shows that this factor leads variations of the policy rate. This reflects the forward-looking information contained in the term structure of LIBOR and futures rates. As market participants anticipate changes in future economic activity and, hence, a tighter or looser policy, *current* interest rates fluctuate. These changes are captured, in part, by the macro factor.

Panel (d) displays the path of the liquidity factor. This factor captures deviations between (risk-adjusted) expectations of Target rates implicit in LIBOR rates relative to those implicit in futures rates. Empirically, the liquidity factor exhibits peaks in periods of tension on financial markets : the Mexican Peso crisis in 1994, the failure of LTCM in 1998 and fears of the Millennium bug. Moreover, the liquidity factor is not unrelated to economic conditions. Figure 2.5 compares the liquidity and the macro factors. These appear to be positively related at the business cycle frequency. In this sample, the contemporaneous correlation of the liquidity factor with the Target rate and the macro factor are 0.42 and 0.38, respectively. However, there are multiple occurrences of negative co-movements at a higher frequency. In many cases, a rapid change in the macro factor is associated with a rapid but opposite change of the liquidity factor. Intuitively, when anticipations of economic conditions change, as revealed by interest rates, the macro factor changes in the same direction (i.e. its loadings are positive). However, the liquidity factor reveals that the spread between futures and LIBOR rates decreases. That is, futures rates increase faster than LIBOR rates when conditions are improving, and decrease faster when conditions are deteriorating and the impact is larger at maturities around 6 months (see LIBOR loadings). In other words, the futures market leads the LIBOR market. Note that this is consistent with a positive price of risk for the liquidity factor. Again, the evidence suggests that the liquidity factor reflects important forward-looking information.

E Pricing Errors

The standard deviations of measurement errors are low, between 0.5 and 10 bps, annually. Mean Pricing Errors [MPE] and Root Mean Squared Pricing Errors [RMSPE] for LIBOR and futures are reported in Panels (a) and (b) of Table III, respectively. Results are reported in bps (annualized). Average pricing errors do not indicate any significant bias, averaging less than 1 bps except for the shortest LIBOR maturities. LIBOR RMSPE decrease with maturities from 6.28 bps at 1 month to below 0.5 bps for the longest maturites. In contrast, futures RMSPE increase with maturity, from 2.56 to 6.18 bps. These are low, by any standard.

VI Forecasting Policy and Excess Returns

Estimates of the structural model completely characterize the distribution of Target rates at any future date via the conditional Laplace transform. Empirically, the model improves on OLS-based forecasts commonly used in practice. Moreover, the improvements increase with the forecast horizon. In fact, I obtain reliable forecasts at horizons up to 12 months. This raises the question as to what factors are the most informative in these forecasts. We know from theory that forecasts based on asset prices rely on the ability to predict excess returns. Therefore, the forecasting results above indicate that the model provides better forecasts of the risk premium. In the following, I measure the relative ability of each factor to explain the risk premium on fed funds futures. As expected, the macroeconomic factor and the Target rate predict variations of the excess returns. However, the liquidity factor is the

Table III: Mean Pricing Errors [MPE] and Root Mean Squared Pricing Errors [RMSPE] using QML estimates. The sample ranges from January 1994 to July 2007. Results are reported in basis points (annualized), for LIBOR rates and futures rates at maturities of 1 to 12 months and 1 to 6 months, respectively.

(a) Mean Pricing Errors

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
LIBOR	-3.17	-1.56	-0.29	-0.11	0.16	0.12	0.10	0.00	-0.23	-0.07	0.04	0.00
Futures	0.28	0.25	-0.06	-0.31	-0.57	-0.76						

(b) Root Mean Square Pricing Errors

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
LIBOR	6.28	4.18	3.28	2.23	1.43	0.72	0.44	0.46	0.61	0.48	0.22	0.54
Futures	2.56	3.24	4.15	5.21	5.92	6.18						

most significant predictor of excess returns. I argue that transitory deviations between LIBOR and futures rates, although due to their relative illiquidity, reflect demand pressure from participants seeking to hedge exposures to future interest rates.

A Forecasting Target Rate

Forecasting Function

In this section, we are interested in forecasting future monthly effective fed funds rate averages, $\bar{r}_{t,n}$. First consider the following forecasting function for the overnight rate at any future date $t+h$,

$$\Psi_{r+s}(x,h) \equiv E_t[r_{t+h} + s_{t+h}],$$

where I use the fact that the information set at time-t can be summarized by the filtered state variables, $x = \hat{X}_t$. Using the the multi-horizon Laplace transform, we have

$$\Psi_{r+s}(x,h) = a_{r+s}(I_{t+1},h) + b_{r+s}(I_{t+1},h)^T x, \tag{2.23}$$

with coefficients $a(\cdot)$ and $b(\cdot)$ given in Appendix D. Then, the time-t forecast of $\bar{r}_{t,n}$ is also linear,

$$\begin{aligned}\Psi_{\bar{r}}(x,n) &\equiv E_t[\bar{r}_{t,n}],\\ &= D_n^{-1}\sum_i a_{r+s}(I_{t+1},h) + b_{r+s}(I_{t+1},h)^T x, &(2.24)\\ &= \bar{a}(I_{t+1},n) + \bar{b}(I_{t+1},n)^T x. &(2.25)\end{aligned}$$

Benchmark Model

As a benchmark of forecasting performances, I use regressions of $\bar{r}_{t,n}$ on current futures rates (Krueger and Kuttner (1996)). Gurkaynak et al. (2007) show that using futures delivers the best market-based forecasts of future policy rates. The forecasting regressions I consider are

$$\begin{aligned}\bar{r}_{t,h} &= \alpha_h^{TS} + \beta_h^{TS}\Psi_r(\hat{X}_t,h) + \epsilon_t^{h,TS}\\ \bar{r}_{t,h} &= \alpha_h^{OLS} + \beta_h^{OLS}F(t,h) + \epsilon_t^{h,OLS}\\ \bar{r}_{t,h}^X &= \alpha_h^X + \beta_h^X\hat{X}_t + \epsilon_t^{h,X},\end{aligned}$$

for each horizon h where $\hat{X}_t$ is the filtered estimates and where $F(t,h)$ is the futures contract for the corresponding months. The first regression uses forecasts from the model. The second regression uses futures rates, and the last regression uses filtered state variables in an unrestricted way. I include this case to see whether the specific combination of state variables imposed by Equation 2.23 uses all the information included in the state variables. The regressions are estimated with daily data for reference months up to 6 months ahead in the case of futures-based regressions and up to 12 months ahead in the case of model-based regressions. In the former case, this implies 184 horizons, and as many regressions, while the latter case implies 360 horizons.

Forecasting Results

The R^2 and RMSE obtained from these regressions measure the forecasting performance. Figure 2.7a presents the R^2 and Figure 2.7b the Root Mean Squared Error [RMSE] in bps (annualized). Forecast errors are typically low at an horizon of 1 month or less, with R^2 very close to 1.[17] This strong result is standard. One-month ahead forecasts of the Target rate are accurate when the information set includes the corresponding futures contract. Forecasting performances then deteriorate slowly as we consider longer horizons. The R^2 of futures-based forecasts decrease to 92% and 78% at 3-month and 6-month horizons, respectively. The corresponding RMSE are 49 bps and 80 bps. This deterioration is halved when we used model-based forecasts. The R^2 from model-based regressions are still 98% at an horizon of 3 months and 92% at an horizon of 6 months. The corresponding RMSE are 23 bps and 47 bps. Moreover, the model extends beyond the longest available contract. At an horizon of one year, model-based forecasts achieve an R^2 and RMSE of 65% and 85 bps, similar to futures-based forecasts at the shorter six month horizon.

Strictly speaking, under the null of the model, Equation 2.23 implies that $\beta_h^{TS} = 1$ and $\alpha_h^{TS} = 0$. This would imply that the model accurately captures the average level and the variations of the risk premium. Figure 2.8a displays coefficient estimates at each horizon. These are close to their theoretical values at each horizon. One year ahead, we have $\hat{\alpha}$ close to 0.25 and and $\hat{\beta}$ close to 0.92. In fact, at these horizons the estimates are not significantly different from 0 and 1, respectively. Panel 2.8b displays the t-statistics of these tests, at each horizon. The tests reject the null only for intermediate

[17] For a one-month forecast horizon, the RMSE is close to 1 bps and 12 bps when using model and futures-based forecasts, respectively. The relative performance of the model at this horizon is due to the implicit adjustment for the remaining number of days. At a 1-day horizon, unadjusted forecasts based on futures are extremely volatile. See Hamilton (2008) for a treatment of the adjustment for the number of days to maturity of the current month contract.

maturities. However, the discrepancies are not economically large, and the statistical evidence disappears at horizons beyond 6 months. The model meets this stringent test with some success. This is a strong and novel result, at these horizons the model provides unbiased forecasts of future monetary policy. Finally, Figure 2.6 includes the R^2 and RMSE from the unrestricted regressions. The results suggest that minor improvements can be achieved at horizons beyond 6 months.

B Forecasting Excess Returns

The model's ability to improve forecasting performances is tightly linked to its ability to predict the risk premium. The realized returns, from entering a futures position today until settlement date is the difference between the current futures rate and the realized monthly average effective rate. That is, $xr_{t,n} = f_t^n - \bar{r}_{t,n}$. Take expectations and re-arrange to get

$$E_t\left[\bar{r}_{t,n}\right] = f_t^n - E_t\left[xr_{t,n}\right],$$

which shows that improvement in forecasting performance, relative to simple futures-based forecasts, must come from the ability to predict excess returns. Then, the forecasting results suggest that the information content of the macroeconomic and liquidity factors is significant. A simple way to evaluate the relative contribution of each state variable is through the following excess returns regressions,

$$xr_{t,h} = \alpha + \beta_r r_t + \beta_s s_t + \beta_z z_t + \beta_l l_t + \epsilon_{t,n}. \tag{2.26}$$

Excess returns from futures contract positions are computed each day, for the 1 to 6-month ahead contract. Results from the regressions are presented in bps (annualized) in Table IV. In-sample, excess returns averaged between 0.66 bps for the one-month contract and up to 15.6 bps for the 6-month contract. Predictability is small at short horizons with R^2 slightly below 1% and 6% for the current month and the month-ahead contract, respectively. However, predictability then rises substantially reaching 12%, 17%, 25% and 27% for the 3, 4, 5 and 6-month ahead contracts. Target rate coefficients are all positive and all significant but for the first month. This implies that periods when the policy rate is high are associated with higher risk premium on futures contracts. A one-standard deviation shock to the Target rate leads, on average, to increases in risk premium of 7 and 19 bps in 3-month and the 6-month contracts, respectively. Next, coefficients on the macroeconomic factor are negative at all horizons, but small and significant only for the intermediate maturities. More important is the information content of the liquidity factor. When futures rates are relatively lower than LIBOR rates, and l_t increases, the risk premia on futures contract also decrease. A one-standard deviation shock to the liquidity factor is associated with a reduction of 6 and 24 basis points for the 3-month and the 6-month contracts, respectively, on average. This effect is large and statistically significant.

Table IV: Excess Returns Regressions

Results from regressions of futures excess returns on state variables obtained from the model,

$$xr_{t+h} = \alpha + \beta_r r_t + \beta_s s_t + \beta_z z_t + \beta_l l_t + \epsilon_{t+h}.$$

Regressors are centered around zero and normalized by their standard deviation and excess returns are in basis points (annualized). Coefficient estimates provide the change in expected excess returns due to a change of one standard deviation in one of the state variables. I include t-statistics based on Newey-West standard errors (30 lags) below each estimate and R^2 at the bottom of each column.

	1	**2**	**3**	**4**	**5**	**6**
α	0.615	2.674	5.194	8.294	12.098	16.369
	(2.217)	(3.631)	(3.956)	(4.220)	(4.616)	(4.882)
β_r	0.430	3.158	6.959	11.109	15.427	19.478
	(0.767)	(2.002)	(2.632)	(2.835)	(2.835)	(2.735)
β_s	-0.019	0.257	-0.043	0.335	-0.129	-0.027
	(-0.094)	(1.227)	(-0.108)	(0.534)	(-0.176)	(-0.032)
β_z	0.187	-0.866	-1.757	-2.089	-0.755	0.889
	(0.277)	(-0.466)	(-0.630)	(-0.517)	(-0.135)	(0.122)
β_l	-0.143	-2.135	-5.573	-10.385	-17.595	-24.285
	(-0.398)	(-2.261)	(-3.318)	(-4.111)	(-5.326)	(-5.849)
R^2	[0.7]	[5.7]	[12.0]	[17.5]	[24.8]	[27.5]

C Discussion

The results above shed light on the interpretation of the liquidity factor. This factor affects model-implied forecasts via three different channels. First, the liquidity factor affects the risk premium associated with Target rate risk through its impact on higher moments of future Target rates. We have that $\hat{\lambda}_{d,l} < \hat{\lambda}_{u,l}$ and this implies that the variance of Target rate changes increases, and skewness is pushed to the right when the liquidity factor decreases. The impact on the risk premium will depends on the relative impact of l_t on the conditional variance and skewness of Target changes. Second, the liquidity factor affects the price of risk of the macro factor directly. Since $\hat{\delta}_{1,l} < 0$ the price of macro risk decreases when the liquidity factor increases. Third, the liquidity factor allows for a wedge between the risk compensation of LIBOR and futures rates. An increase in the liquidity factor may be caused by an increase of LIBOR rates or a decrease of futures rates. The latter case induces lower excess returns on futures following a shock to the liquidity factor. These last two channels imply negative liquidity coefficients in excess return regressions.

Empirically, increases in the liquidity factor predicts lower excess returns on futures contracts and conditioning for the forward-looking information contained in the liquidity premium is key to improve forecasting performances. In other word, when the liquidity factor increases, futures rates tend to be lower on average relative to LIBOR rates and relative to future realized monthly average. This suggests that the information content of the liquidity factor should be interpreted in terms of futures rates relative to LIBOR rates. A likely explanation is that hedging demand put pressures on the intermediation mechanisms of futures markets and causes transitory deviations of futures rate to compensate liquidity providers for their services. While I assumed that only the LIBOR market suffered from illiquidity, it is likely that both markets are affected by intermediation frictions, but to different degrees. That shocks to hedging demand are revealed in futures markets suggests that they are more liquid, less costly, or that participation is less limited, than LIBOR markets. This complements the evidence provided in Piazzesi and Swanson (2006) who showed that excess returns on futures varies with banks' hedging demand in anticipation of policy changes.

Finally, note that the lack of a perfectly liquid interest rate contract and of measures of the frictions interest rate markets imply that we cannot identify the liquidity component in each market separately. However, we can measure their difference. This is precisely what estimates of l_T achieve given the assumption of a zero liquidity component in futures markets. In particular, the liquidity component identified here cannot be directly measured from the spread between LIBOR and futures rates.

VII Conclusion

This papers provides a joint model of the monetary policy response function, of LIBOR rates and of Fed funds futures rates. Combined with no-arbitrage restrictions, this provides better identification of the price of risk parameters and delivers significant improvements in policy rate forecasts. Moreover, and perhaps unexpectedly, allowing for transitory deviations between these two markets captures forward-looking information about the future path of the policy rate. Demand for immediacy pressures the intermediation mechanism of futures markets. Note that this liquidity premium is hard to disentangle from other components of futures rates without a joint model. In any case, the evidence leads to the important conclusion that liquidity premium in interest rate markets are relevant to macroeconomic researchers.

VIII Appendix

A Conditional Laplace Transform

Skellam Distribution

The dynamics of state variables is summarized by Equations 2.1, 2.5, 2.6 and 2.7. Although the process for r_t is novel, the conditional multivariate Laplace Transform of the state variables is exponential affine. Conditional on the past, changes in the Target rate follow a Skellam distribution (Skellam (1946), Johnson et al. (1997)). This distribution is characterized as the difference between two independent Poisson random variables. Both its Laplace transform and its density are known in closed-form. Consider two univariate Poisson variables N_1 and N_2 with parameters λ_1 and λ_2, respectively. The Laplace Transform of their difference $Z \equiv N_1 - N_2$ is

$$T(u, Z) = \exp\left(\lambda_1(e^u - 1) + \lambda_2(e^{-u} - 1)\right), \quad u \in R,$$

while its probability mass function is

$$f(Z = z) = \exp(-(\lambda_1 + \lambda_2)) \left(\frac{\lambda_1}{\lambda_2}\right)^{z/2} I_z\left(2\sqrt{\lambda_1\lambda_2}\right),$$

where $I_k(y)$ is the modified Bessel function of the first kind. In our context, the coefficients of the conditional Laplace transform vary through time because of the evolution of the underlying state variables and because of the (deterministic) variation in the FOMC meeting schedule. Hence, computation of the transform depends on the occurrence of a FOMC meeting in the next period and the two cases must treated separately.

No FOMC Meeting

I first consider the case where no meeting is scheduled to occur. Then, since the innovations driving the state vector, X_t, are independent, the conditional transform is

$$\begin{aligned} T(u, X_{t+1}|I_{t+1} = 0) &\equiv E_t\left[\exp\left(u^T X_{t+1}\right)|I_{t+1} = 0\right] \\ &= \exp\left(A(I_{t+1} = 0, u) + B(I_{t+1} = 0, u)^T X_t\right), \end{aligned}$$

where I_t is a binary variable that is equal to 1 if a meeting occurs at time-t and 0 otherwise. The coefficients are given by

$$\begin{aligned} A(I_t = 0, u) &= \lambda_0^u(e^{\Delta u_r} - 1) + \lambda_0^d(e^{-\Delta u_r} - 1) \\ &+ u_s\mu_s + u_z\mu_z + u_l\mu_l + \frac{1}{2}(u_s^2\sigma_s^2 + u_z^2\sigma_z^2 + u_l^2\sigma_l^2) \\ &+ \lambda_{J,s}\left(T(u_s, J_{t+1}^s) - 1\right) \\ &= g_0(u_r) + u^T\mu + \frac{1}{2}u^T\Omega u + \lambda_{J,s}\left(T(u_s, J_{t+1}^s) - 1\right) \end{aligned}$$

and

$$B(I_t = 0, z) = [u_r \quad u_s\phi_s \quad u_z\phi_z \quad u_l\phi_l]^T$$
$$= \Phi u,$$

where I defined

$$\mu = [0 \quad \mu_s \quad \mu_z \quad \mu_l]^T$$
$$\Phi = diag([1 \quad \phi_s \quad \phi_z \quad \phi_l]^T)$$
$$\Omega = diag([0 \quad \sigma_s^2 \quad \sigma_z^2 \quad \sigma_l^2]^T),$$

$g_0(x) = \lambda_0^u(e^{\Delta u_r} - 1) + \lambda_0^d(e^{-\Delta u_r} - 1)$ and where $diag$ is the usual operator from the vector space to the space of diagonal matrices.

FOMC Meeting

Computation of the multivariate conditional Laplace transform is slightly different when a FOMC meeting occurs in the following period. Conditional on the realization of s_{t+1}, z_{t+1} and l_{t+1}, we have

$$T(u, X_{t+1}|I_{t+1} = 1) \equiv E_t\left[\exp\left(u^T X_{t+1}\right)|I_{t+1} = 1\right]$$
$$= E_t\left[E_{s_{t+1},z_{t+1},l_{t+1}}\left[\exp\left(u^T X_{t+1}\right)|I_{t+1} = 1\right]\right]$$
$$= E_t\left[\exp\left(-G(u_r)^T\bar{X} + \lambda(e^{\Delta u_r} + e^{-\Delta u_r} - 2) + G(u_r)^T X_{t+1}^* + u^T X_{t+1}^*\right)\right],$$

where

$$X_{t+1}^* = [r_t \; s_{t+1} \; z_{t+1} \; l_{t+1}]$$
$$G(y) = [g_r(y) \; g_r(y) \; g_z(y) \; g_l(y)]^T, \; y \in R,$$

and the functions $g_k(y), k = r, s, z,$ are defined as

$$g_k(y) \equiv \left(\lambda_{u,k}(e^{y\Delta} - 1) - \lambda_{d,k}(e^{-y\Delta} - 1)\right).$$

We then have that

$$T(u, X_{t+1}|I_{t+1} = 1) = \exp\left(A(I_t = 1, u) + B(I_t = 1, u)^T X_t\right),$$

with coefficients

$$A(1, u) = -G(u_r)^T\bar{X} + \lambda(e^{\Delta u_r} - 1) + \lambda(e^{-\Delta u_r} - 1) + k_s(u_r)\mu_s + k_z(u_r)\mu_z + k_l(u_r)\mu_l$$
$$+ \frac{1}{2}(k_s(u_r)^2\sigma_s^2 + k_z(u_r)^2\sigma_z^2 + k_l(u_r)^2\sigma_l^2)$$
$$+ \lambda_{J,s}\left(T(k_s(u_r), J_{t+1}^s) - 1\right)$$
$$= -G(u_r)^T\bar{X} + h_1(u_r) + (G(u_r) + u)^T\mu + \frac{1}{2}(G(u_r) + u)^T\Omega(G(u_r) + u)$$
$$+\lambda_{J,s}\left(T(g_s(u_r) + u_s, J_{t+1}^s) - 1\right)$$

and

$$
\begin{aligned}
B(1,u) =& [g_r(u_r)+u_r \quad (g_s(u_r)+u_s)\phi_s \quad (g_z(u_r)+u_z)\phi_z \quad (g_l(u_r)+u_l)\phi_l]^T \\
=& \Phi\left(G(u_r)+u\right).
\end{aligned}
$$

Risk-Neutral Distribution

It is easy to show using results from the previous sections that

$$
\begin{aligned}
T^Q(u,X_{t+1}) &= E_t\left[M_{t,t+1}\exp\left(u^T X_{t+1}\right)\right] \\
&= \exp\left(A^Q(I_{t+1},u)+(B^Q(I_{t+1},u)-C_r)^T X_t\right),
\end{aligned}
$$

with $C_r = [1\ 1\ 0\ 0]^T$ and coefficients given by

$$
\begin{aligned}
A^Q(0,u) =& g_0(u_r+\delta_{0,r})-g_0(\delta_{0,r})+u^T(\mu+\Sigma\delta_0)+\frac{1}{2}u^T\Omega u \\
&+\lambda_{J,s}\left(T(u_s+\delta_{0,s},J^s_{t+1})-T(\delta_{0,s},J^s_{t+1})\right) \\
A^Q(1,u) =& g_1(u_r+\delta_{0,r})-g_1(\delta_{0,r})-\left(G(u_r+\delta_{0,r})-G(\delta_{0,r})\right)^T\bar{X} \\
&-\frac{1}{2}G(\delta_{0,r})^T\Omega G(\delta_{0,r})+\frac{1}{2}\left(G(u_r+\delta_{0,r})+u\right)^T\Omega\left(G(u_r+\delta_{0,r})+u\right) \\
&+\left(G(u_r+\delta_{0,r})+u\right)^T(\mu+\Omega\delta_0) \\
&+\lambda_{J,s}\left(T(g_s(u_r+\delta_{0,r})+\delta_{0,s}+u_s,J^s_{t+1})-T(g_s(\delta_{0,r})+\delta_{0,s},J^s_{t+1})\right)
\end{aligned}
$$

and

$$
\begin{aligned}
B^Q(0,u) &= (\Phi+\Omega\delta_1)\,u \\
B^Q(1,u) &= (\Phi+\Omega\delta_1)\left(G(u_r+\delta_{0,r})-G(\delta_{0,r})+u\right).
\end{aligned}
$$

Note that the persistence of the factors is shifted whenever $\delta_0 \neq 0$. That is, the persistence under the risk-neutral measures will be shifted whenever interest rate risk is priced even if the vector of prices of risk is constant (i.e. $\delta_1 = 0$).

B Generating Function for Prices

Consider the price at time-t of the payoff $\exp(u^T X_{t+m})$ at maturity m,

$$
\begin{aligned}
\Gamma(u,t,m) &= E_t\left[M_{t,t+m}\exp\left(u^T X_{t+m}\right)\right] \\
&= E_t\left[M_{t+1}\Gamma\left(u,t+1,m-1\right)\right].
\end{aligned}
$$

Substituting the guess $\Gamma(u,t,m)=\exp\left(c_0(u,t,m)+c(u,t,m)^T X_t\right)$ gives

$$
\begin{aligned}
\Gamma(u,t,m) =& E_t\left[M_{t+1}\exp\left(c_0(u,t+1,m-1)+c(u,t+1,m-1)^T X_{t+1}\right)\right] \\
=& \exp\left(c_0(u,t+1,m-1)+A^Q(I_{t+1},c(u,t+1,m-1))\right) \\
&\times\exp\left(\left[B^Q(I_{t+1},c(u,t+1,m-1))-C_r\right]^T X_t\right),
\end{aligned}
$$

which implies the following recursion that coefficients must solve:

$$c_0(u,t,h) = c_0(u,t+1,h-1) + A^Q(I_{t+1}, c(u,t+1,h-1)) \tag{2.27}$$
$$c(u,t,h) = B^Q(I_{t+1}, c(u,t+1,h-1)) - C_r, \tag{2.28}$$

for $0 \leq h \leq m$. Note that $\Gamma(u,t,0) = \exp(u^T X_t)$ implies $c_0(u,t,0) = 0$ and $c(u,t,0) = u$ for arbitrary $t \geq 1$ and u.

C Asset Prices

Discount Bonds

Of particular interest is the case $u = 0$. This corresponds to the price, $D(t,m)$, of a risk-free discount bond with maturity m,

$$\begin{aligned} D^{rf}(t,m) &\equiv \Gamma(0,t,m) = E_t[M_{t,t+m}] \\ &= \exp\left(d_0^{rf}(t,m) + d^{rf}(t,m)^T X_t\right), \end{aligned}$$

where $d_0^{rf}(t,m) \equiv c_0(0,t,m)$ and $d^{rf}(t,m) \equiv c(0,t,m)$.

LIBOR loan

A LIBOR loan is an asset with unit payoff which is further discounted at the rate l_t to offer compensation for illiquidity or counterparty risk. The price of a LIBOR loan can also be obtained from the price generating function by noting that

$$\begin{aligned} D^L(t,m) &= E_t\left[M_{t,t+m} \exp\left(-\sum_{i=0}^{m-1} l_{t+i}\right)\right] \\ &= E_t\left[M^L_{t,t+m}\right] \end{aligned}$$

with $M^L_{t+i} = M_{t+i}\exp(-l_{t+1})$. I guess and verify that the solution is exponential-affine, $D^L(t,m) = \exp\left(d_0^L(t,m) + d^L(t,m)^T X_t\right)$ with solution

$$\begin{aligned} d_0^L(t,m) &= d_0^L(t+1,m-1) + A^Q\left(I_{t+1}, d^L(t+1,m-1)\right) \\ d^L(t,m) &= B^Q\left(I_{t+1}, d^L(t+1,m-1)\right) - C_L, \end{aligned}$$

where $C_L = [1\ 1\ 0\ 1]^T$. Finally, note that $D^L(t,0) = 1$ implies that $d_0^L(t,0) = 0$ and $d^L(t,0) = 0$ for any $t \geq 1$.

Singleton futures Price

A difficulty arises when computing a singleton futures rate because the reference date $t+m$, at which the payoff is determined, is not generally the same than the settlement date $t+T$, at which the payment is made. That is, we have

$$\begin{aligned} f(t,m,T) &= E_t[M_{t,t+T} r_{t+m}] = \\ &= \frac{\partial}{\partial u} E_t\left[M_{t,T}\exp(u r_{t+m})\right]\Big|_{u=0} \equiv \frac{\partial}{\partial u}\Gamma_f(u,t,m,T)\Big|_{u=0}, \end{aligned}$$

where $m \leq T$ and $u \in R$. It seems at first that the price generating function derived above will not help whenever m is different than T. However, we can use the law of iterated expectations and get

$$\begin{aligned}\Gamma_f(u,t,m,T) &= E_t\left[M_{t,t+m}\exp(ur_{t+m})E_{t+m}\left[M_{t+m,t+T}\right]\right] \\ &= E_t\left[M_{t,t+m}\exp(ur_{t+m})D(t+m,T-m)\right] \\ &= E_t\left[M_{t,t+m}\exp\left(d_0^{rf}(t+m,T-m)+[d^{rf}(t+m,T-m)+uC_r]^T\tilde{X}_{t+h}\right)\right] \\ &= \exp(d_0^{rf}(t+m,T-m))\Gamma(d^{rf}(t+m,T-m)+uC_r,t,m).\end{aligned}$$

We can then use the results above to obtain

$$\begin{aligned}&\Gamma_f(u,t,m,T) = \exp\left(d_0^{rf}(t+m,T-m)\right) \\ &\times \exp\left(c_0(d^{rf}(t+m,T-m)+uC_r,t,m)+c(d^{rf}(t+m,T-m)+uC_r,t,m)^T X_t\right).\end{aligned}$$

Taking the partial derivatives with respect to u and evaluating at $u=0$, the singleton futures rate is

$$\begin{aligned}f(t,m,T) = \exp&\left(d_0^{rf}(t+m,T-m)+c_0\left(u^*,t,m\right)+c\left(u^*,t,m\right)^T X_t\right) \\ &\times \left[c_0'\left(u^*,t,m\right)+X_t^T c'\left(u^*,t,m\right)\right]C_r,\end{aligned}$$

where $u^* = d^{rf}(t+m,T-m)$. Note that u^* is only a function of the reference date for the singleton futures, $t+h$, and the length of time between the reference date and the settlement date, which will not change as we vary t or m in the coefficient recursions. That is, for a given set of risk-free zero coupon coefficients, we can apply the same strategy as for simple interest rates to compute futures coefficients. The differentiated coefficients, $c_0'(\cdot)$ and $c'(\cdot)$, can be computed by taking derivatives with respect to u on both sides of Equation (2.27). They must satisfy:

$$\begin{aligned}c_0'(u,t,h) &= c_0'(u,t+1,h-1)+A'^Q\left(I_{t+1},c(u,t+1,h-1)+\delta\right)c'\left(u,t+1,h-1\right) \\ c'(u,t,h) &= B'^Q\left(I_{t+1},c(u,t+1,h-1)\right)c'(u,t+1,h-1)\end{aligned}$$

for any u, t and any $h>0$. Initial conditions for these differentiated recursions can be found by differentiation of the corresponding initial conditions or by nothing that we must have $f(t,0,T) = D(t,T)r_t$. This yields $c_0'\left(u,t,0\right)=0$ and $c'\left(u,t,0\right)=Id$, where Id is the identity matrix.

Finally, the derivatives of Laplace coefficients, $A^{Q'}(\cdot)$ and $B^{Q'}(\cdot)$ can be computed directly,

$$A'^Q(0,u) = \begin{bmatrix} g_0'(u_r+\delta_{0,r}) \\ \mu_s+(u_s+\delta_s)\sigma_s^2+\lambda_s T\left((u_s+\delta_s),J_{t+1}^s\right)\left(\nu_s+(u_s+\delta_s)\omega_s^2\right) \\ \mu_z+(u_z+\delta_z)\sigma_z^2 \\ \mu_l+(u_l+\delta_l)\sigma_l^2\end{bmatrix}^T,$$

$$
\begin{aligned}
A'^{Q}(1,u)_1 =& g_1'(u_r+\delta_r) - G'_{(r)}(u_r+\delta_r)^T\bar{X} + G'_{(r)}(u_r+\delta_r)^T(\mu+\Omega\delta_0)\\
&+\lambda_s T\left(g_s(u_r+\delta_r)+u_s+\delta_s, J^s_{t+1}\right)\left(\nu_s+(g_s(u_r+\delta_r)+u_s+\delta_s)\omega_s^2\right)g_s'(u_r+\delta_r))\\
&+\sum_{i=z,k,l}\left(g_i(u_r+\delta_r)g_i'(u_r+\delta_r)\sigma_i^2\right)\\
A'^{Q}(1,u)_2 =& \mu_s+\delta_s\sigma_s^2+(g_s(u_r+\delta_{0,r})+u_s)\,\sigma_s^2\\
&+\lambda_s T\left(g_s(u_r+\delta_r)+u_s+\delta_s, J^s_{t+1}\right)\left(\nu_s+(g_s(u_r+\delta_r)+u_s+\delta_s)\omega_s^2\right)\\
A'^{Q}(1,u)_3 =& \mu_z+\delta_z\sigma_z^2+(g_s(u_z+\delta_{0,r})+u_z)\,\sigma_z^2\\
A'^{Q}(1,u)_4 =& \mu_l+\delta_l\sigma_l^2+(g_s(u_l+\delta_{0,r})+u_l)\,\sigma_l^2,
\end{aligned}
$$

$$
B'^{Q}(0,u) = [\Phi+\Omega\delta_1],
$$

$$
B'^{Q}(1,u) = \begin{bmatrix}
\sum_{i=r,s,z,l} g_i'(u_r+\delta_{0,r})(\phi_{ri}+\delta_{0,ir}\sigma_i^2) & 0 & \sigma_z^2\delta_{i,zr} & \sigma_l^2\delta_{i,lr}\\
\sum_{i=r,s,z,l} g_i'(u_s+\delta_{0,s})(\phi_{si}+\delta_{0,is}\sigma_i^2) & \phi_s & \sigma_z^2\delta_{i,zs} & \sigma_l^2\delta_{i,ls}\\
\sum_{i=r,s,z,l} g_i'(u_z+\delta_{0,z})(\phi_{zi}+\delta_{0,iz}\sigma_i^2) & 0 & \phi_z+\sigma_z^2\delta_{i,zz} & \sigma_l^2\delta_{i,lz}\\
\sum_{i=r,s,z,l} g_i'(u_l+\delta_{0,l})(\phi_{li}+\delta_{0,il}\sigma_i^2) & 0 & \sigma_z^2\delta_{i,zl} & \phi_l+\sigma_l^2\delta_{i,ll}
\end{bmatrix}.
$$

D Predictability Coefficients

Multi-Horizon Laplace Transform

The distribution of future state variables can be characterized explicitly from the multi-horizon conditional Laplace transform,

$$
T_X(u,t,h) \equiv E_t\left[\exp(u^T X_{t+h})\right],
$$

for any $u \in R^K$ and $h \geq 1$. This can be derived from the known one-horizon case, guessing an exponential affine solution and noting that

$$
\begin{aligned}
T_X(u,t,h) \equiv& E_t\left[T_X(u,t+1,h-1)\right],\\
=& \exp\left(A(I_{t+1},z,h)+B(I_{t+1},z,h)^T X_t\right),
\end{aligned}
$$

with coefficients given by

$$
\begin{aligned}
A(I_t,z,h+1) =& A(I_{t+1},z,h)+A(I_t,B(I_{t+1},z,h))\\
B(I_t,z,h+1) =& B(I_t,B(I_{t+1},z,h)),
\end{aligned}
$$

for any $h \geq 1$ with initial conditions $A(I_{t+1},z,1)=A(I_{t+1},z)$ and $B(I_{t+1},z,1)=B(I_{t+1},z)$.

Forecast Functions

The forecast function for any linear combination, $C^T X_t$ of the state variables can be derived at any horizon from the following partial derivative with respect to u,

$$
\begin{aligned}
E_t\left[C^T X_{t+h}\right] &= \left[\frac{\partial}{\partial u}T_X(uC,t,h)\right]_{u=0}\\
&= \left[T_X(uC,t,h)\left(A_2'(I_{t+1},uC,h)+X_t^T B_2'(I_{t+1},uC,h))\right)C\right]_{u=0}\\
&= \left(A_2'(I_{t+1},0,h)+X_t^T B_2'(I_{t+1},0,h))\right)C,
\end{aligned}
$$

where, as before, the derivatives of the multi-horizon coefficients can be obtained by differentiating their respective recursions. The derivatives of the Laplace Transform coefficients under the historical measure are given by

$$A'(I_t = 0, u)^T = \begin{bmatrix} \Delta\lambda_0(\exp(\Delta u_r) - \exp(-\Delta u_r)) \\ \mu_s + u_s\sigma_s^2 + \lambda_s T(u_s, J_{t+1}^s)(\nu_s + u_s\omega_s^2) \\ \mu_z + u_z\sigma_z^2 \\ \mu_l + u_l\sigma_l^2 \end{bmatrix},$$

$$A'(I_t = 1, u)^T = \begin{bmatrix} -G(u_r)^T\bar{X} + \lambda\Delta(e^{\Delta u_r} - e^{-\Delta u_r}) + G'(u_r)\mu + {}^{\backprime}u_s\sigma_s^2 + u_z\sigma_z^2 + u_3\sigma_r^2 + \lambda_s T(u_s, J_{t+}^s \\ u_s + u_s(u_r)\sigma_s^2 + \lambda_s T\left(k_s(u_r), J_{t+1}^s\right)\left(\mu_s + k_s(u_r)\sigma_s^2\right) \\ u_z + k_z(u_r)\sigma_z^2 \\ u_l + k_l(u_r)\sigma_l^2 \end{bmatrix}$$

$$B'(I_t = 0, u) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \phi_s & 0 & 0 \\ 0 & 0 & \phi_z & 0 \\ 0 & 0 & 0 & \phi_l \end{bmatrix},$$

$$B'(I_t = 1, u) = \begin{bmatrix} g_r'(u_r) + 1 & 0 & 0 & 0 \\ g_s'(u_r)\phi_s & \phi_s & 0 & 0 \\ g_x'(u_r)\phi_z & 0 & \phi_z & 0 \\ g_l'(u_r)\phi_l & 0 & & \phi_r \end{bmatrix},$$

and, hence, the multi-horizon derivatives with respect to the second argument are

$$\begin{aligned} A_2'(I_{t+1}, 0, h) =& A_2'(I_{t+2}, 0, h-1) + A_2'(I_{t+1}, B(I_{t+2}, 0, h-1))B_2'(I_{t+2}, 0, h-1)) \\ B_2'(I_{t+1}, 0, h)) =& B_2'(I_{t+2}, B(I_{t+2}, 0, h-1))B_2'(I_{t+2}, 0, h-1)), \end{aligned}$$

with initial conditions $A_2'(I_{t+1}, 0, 1) = A_2'(I_{t+1}, 0)$ and $B_2'(I_{t+1}, 0, 1) = B_2'(I_{t+1}, 0)$.

Target and Effective Forecast

The forecast function for future Target and Effective overnight Fed funds rates can by computed by setting $C = C_r = [1\ 0\ 0\ 0]^T$ and $C = C_{r+s} = [1\ 1\ 0\ 0]^T$, respectively. We then have

$$\begin{aligned} \Psi_r(x, t, h) =& E[r_{t+h}|X_t = x] \\ =& a_r(I_{t+1}, h) + b_r(I_{t+1}, h)^T X_t \\ \Psi_{r+s}(x, t, h) =& E[r_{t+h}|X_t = x] \\ =& a_{r+s}(I_{t+1}, h) + b_{r+s}(I_{t+1}, h)^T X_t. \end{aligned}$$

E *Unscented Kalman Filter*

The UKF is based on an approximation to any non-linear transformation of a probability distribution. It has been introduced in Julier et al. (1995) and Julier and Uhlmann (1996) (see Wan and der Merwe (2001) for textbook treatment) and was first imported in finance by Leippold and Wu (2003).

Given $\hat{X}_{t+1|t}$ a time-t forecast of state variable for period $t+1$, and its associated MSE $\hat{Q}_{t+1|t}$ the unscented filter selects a set of Sigma points in the distribution of $X_{t+1|t}$ such that

$$\bar{\mathbf{x}} = \sum_i w^{(i)} x^{(i)} = \hat{X}_{t+1|t}$$
$$\mathbf{Q}_x = \sum_i w^{(i)} (x^{(i)} - \bar{\mathbf{x}})(x^{(i)} - \bar{\mathbf{x}})' = \hat{Q}_{t+1|t}.$$

Julier et al. (1995) proposed the following set of Sigma points,

$$x^{(i)} = \begin{cases} \bar{\mathbf{x}} & i = 0 \\ \bar{\mathbf{x}} + \left(\sqrt{\frac{N_x}{1-w^{(0)}} \sum_x}\right)_{(i)} & i = 1, \ldots, K \\ \bar{\mathbf{x}} - \left(\sqrt{\frac{N_x}{1-w^{(0)}} \sum_x}\right)_{(i-K)} & i = K+1, \ldots, 2K \end{cases}$$

with weights

$$w^{(i)} = \begin{cases} w^{(0)} & i = 0 \\ \frac{1-w^{(0)}}{2K} & i = 1, \ldots, K \\ \frac{1-w^{(0)}}{2K} & i = K+1, \ldots, 2K \end{cases}$$

where $\left(\sqrt{\frac{N_x}{1-w^{(0)}} \sum_x}\right)_{(i)}$ is the i-th row or column of the matrix square root. Julier and Uhlmann (1996) use a Taylor expansion to evaluate the approximation's accuracy. The expansion of $y = g(x)$ around $\bar{x}$ is

$$\begin{aligned} \bar{y} &= E\left[g(\bar{x} + \Delta x)\right] \\ &= g(\bar{x}) + E\left[D_{\Delta x}(g) + \frac{D^2_{\Delta x}(g)}{2!} + \frac{D^3_{\Delta x}(g)}{3!} + \cdots\right], \end{aligned}$$

where the $D^i_{\Delta x}(g)$ operator evaluates the total differential of $g(\cdot)$ when perturbed by Δx, and evaluated at $\bar{x}$. A useful representation of this operator in our context is

$$\frac{D^i_{\Delta x}(g)}{i!} = \frac{1}{i!} \left(\sum_{j=1}^{n} \Delta x_j \frac{\partial}{\partial x_j} \right)^i g(x) \Bigg|_{x=\bar{x}}.$$

Different approximation strategies for $\bar{y}$ will differ by either the number of terms used in the expansion or the set of perturbations Δx. If the distribution of Δx is symmetric, all odd-ordered terms are zero. Moreover, we can re-write the second terms as a function of the covariance matrix P_{xx} of Δx,

$$\bar{y} = g(\bar{x}) + \left(\nabla^T P_{xx} \nabla\right) g(\bar{x}) + E\left[\frac{D^4_{\Delta x}(g)}{4!} + \cdots\right].$$

Linearisation leads to the approximation $\hat{\bar{y}}_{lin} = g(\bar{x})$ while the unscented approximation is exact up to the third-order term and the σ-points have the correct covariance matrix by construction. In the Gaussian case, Julier and Uhlmann (1996) show that same-variable fourth moments agree as well and that all other moments are lower than the true moments of Δx. Then, approximation errors of higher order terms are necessarily smaller for the UKF than for the EKF. Using a similar argument to the case of the approximation of the MSE, Julier and Uhlmann (1996) show that linearization and the unscented transformation agree with the Taylor expansion up to the second-order term and that approximation errors in higher-order terms are smaller for the UKF.

Figure 2.2: Time series of forward rates and futures rates at maturities of 2, 4 and 6 months. Forward rates are computed from LIBOR rates.

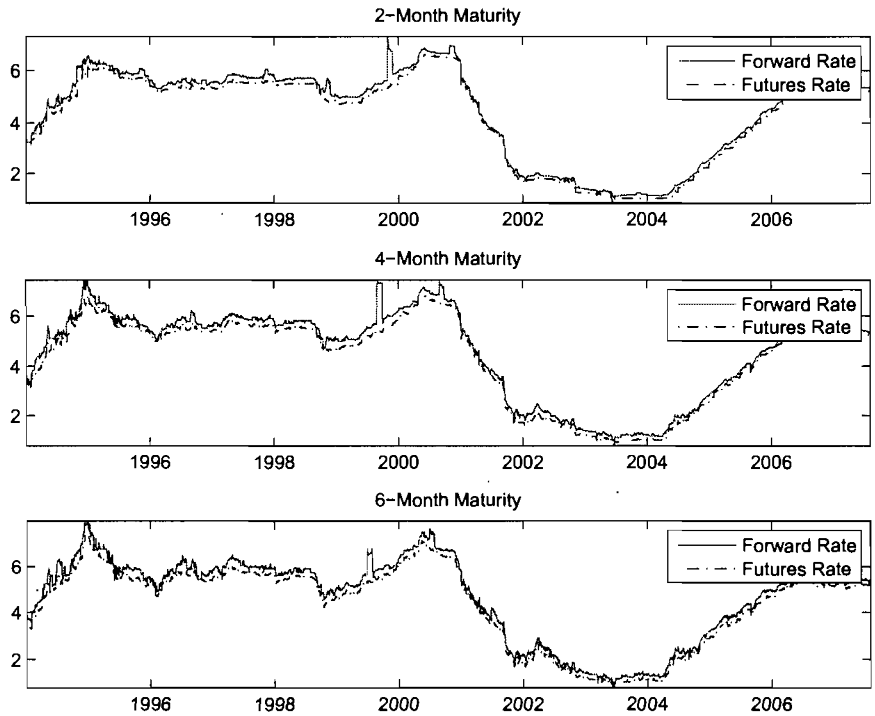
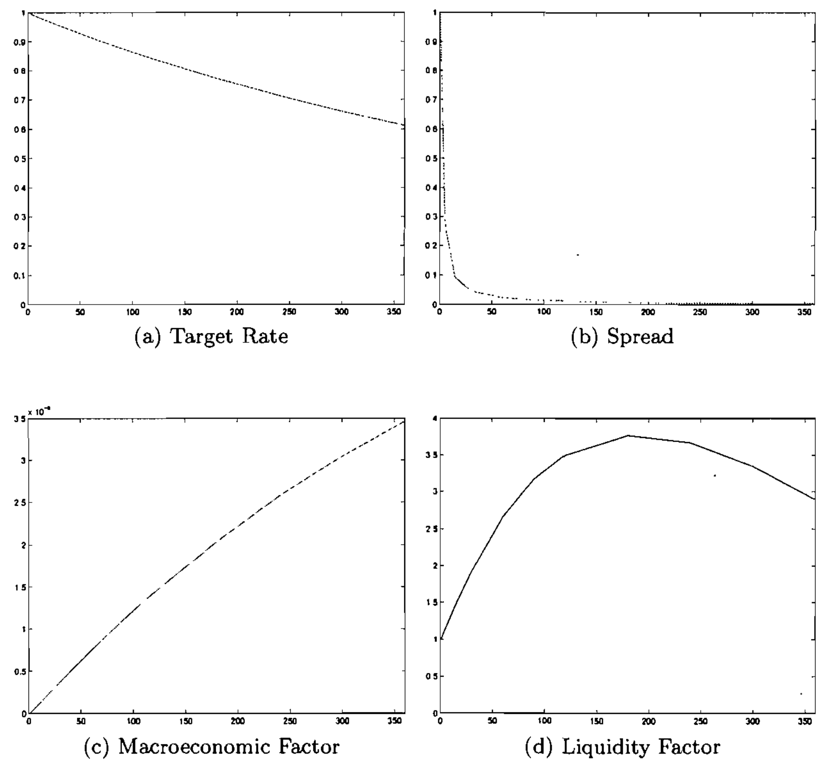


Figure 2.3: Libor rate factor loadings computed from the model. Panel 2.4a presents loadings on the Target rate. Panel 2.4b presents loadings on the effective spread. Panel 2.4c presents loadings on the macro factor. Panel 2.4d presents loadings on the liquidity factor.



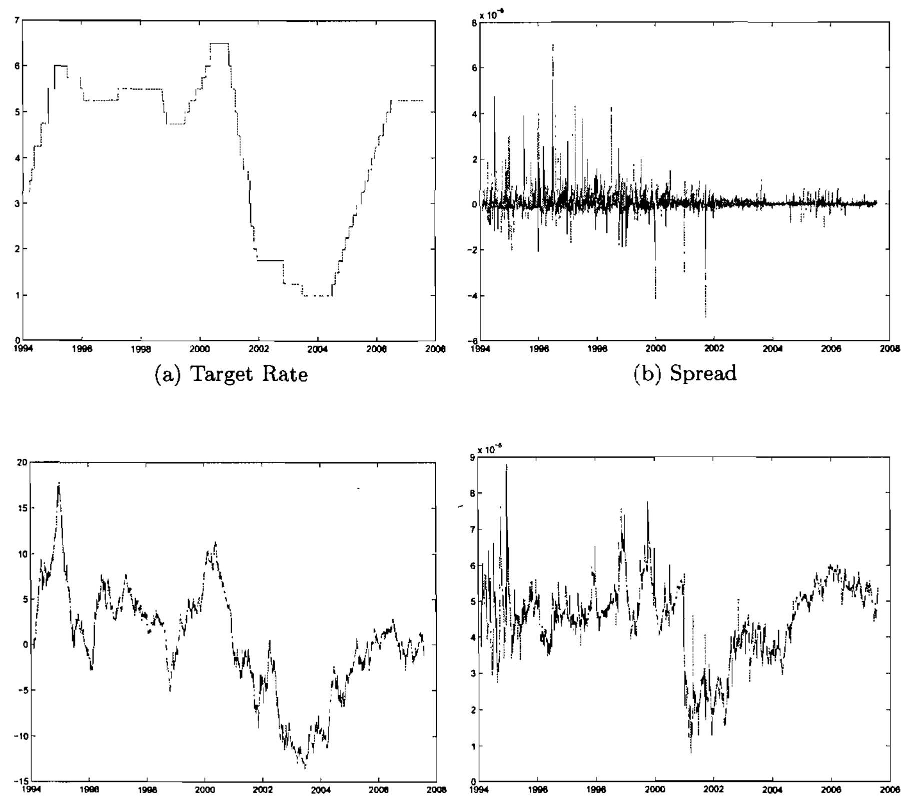
(a) Target Rate

(b) Spread

(c) Macroeconomic Factor

(d) Liquidity Factor

Figure 2.4: Filtered state variables from QML estimation of the model in a a daily sample of Target, effective, LIBOR and futures rates January 1994 to July 2007. Panel 2.5a displays the time series of the Target rate. Panel 2.5b displays the time series of the effective spread. Panel 2.5c displays the time series of the macro factor. Panel 2.5d displays the time series of the liquidity factor.



(a) Target Rate

(b) Spread

(c) Macroeconomic Factor

(d) Liquidity Factor

Figure 2.5: Filtered time-series of the Liquidity and Macroeconomic and factors from the unrestricted model. Factors are reported with standardized units from Jan. 1st 1994 to the end of July 2007.

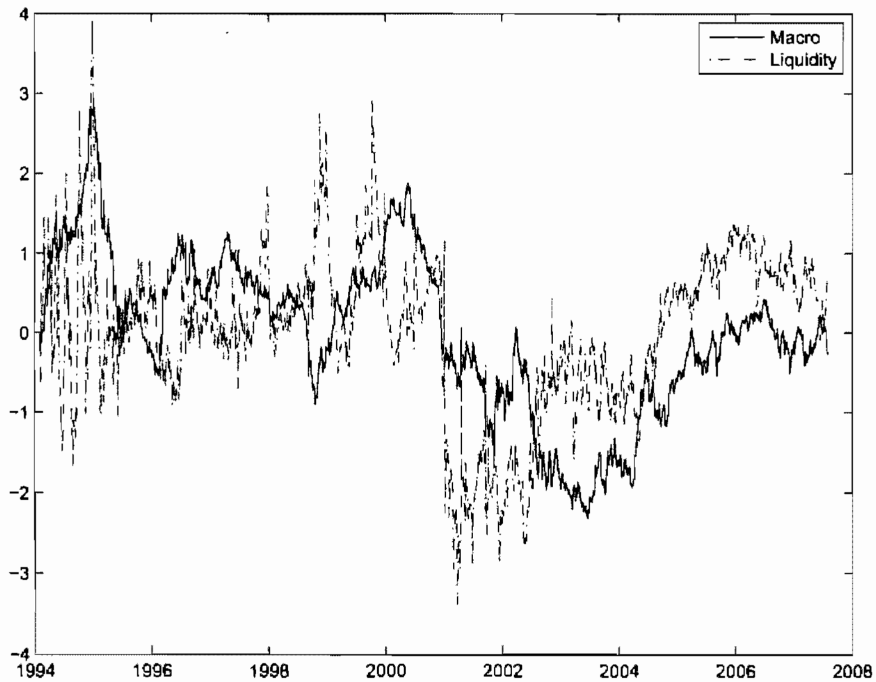
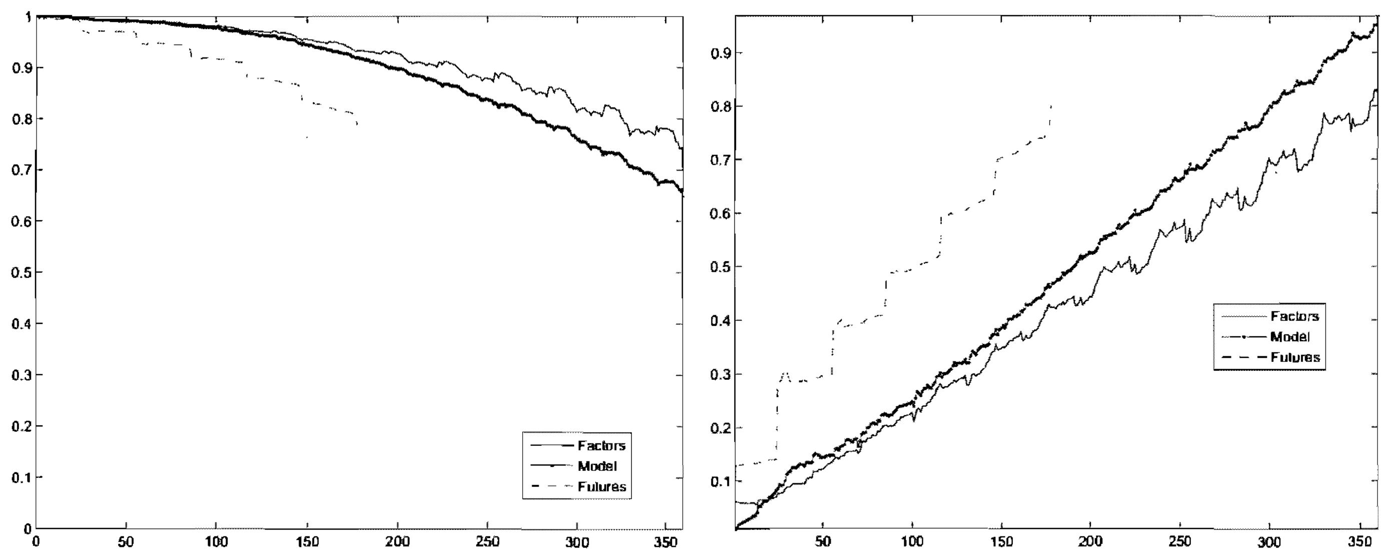


Figure 2.6: R2 and RMSE from forecasting regressions

Comparing forecasting performance of model-forecast based on the following regressions of realized monthly average Target rate, $\bar{r}_{t,n}^{TS}$,

$$\bar{r}_{t,n} = \alpha_n^{TS} + \beta_n^{TS}\Psi_{\bar{r}}(\hat{X}_t, n) + \epsilon_t^{n,TS}$$
$$\bar{r}_{t,n} = \alpha_n^{Fut} + \beta_n^{OLS} f(t,n) + \epsilon_t^{n,OLS}$$
$$\bar{r}_{t+n} = \alpha_n^{X} + \beta_n^{X}\hat{X}_t + \epsilon_t^{n,X}$$

where $\Psi_r(\hat{X}_t, n)$ is the model's forecast, $\hat{X}_t$ are filtered state variables from the model and $f(t,n)$ is the observed futures prices corresponding to a time-t, n-month ahead, forecast of the monthly average of the effective Fed funds rate. Panel (a) compares the R^2 from each regression (y-axis is from 0 to 1) and Panel (b) compares the RMSE in bps (annualized) from each regression.



(a) R^2 from each regression

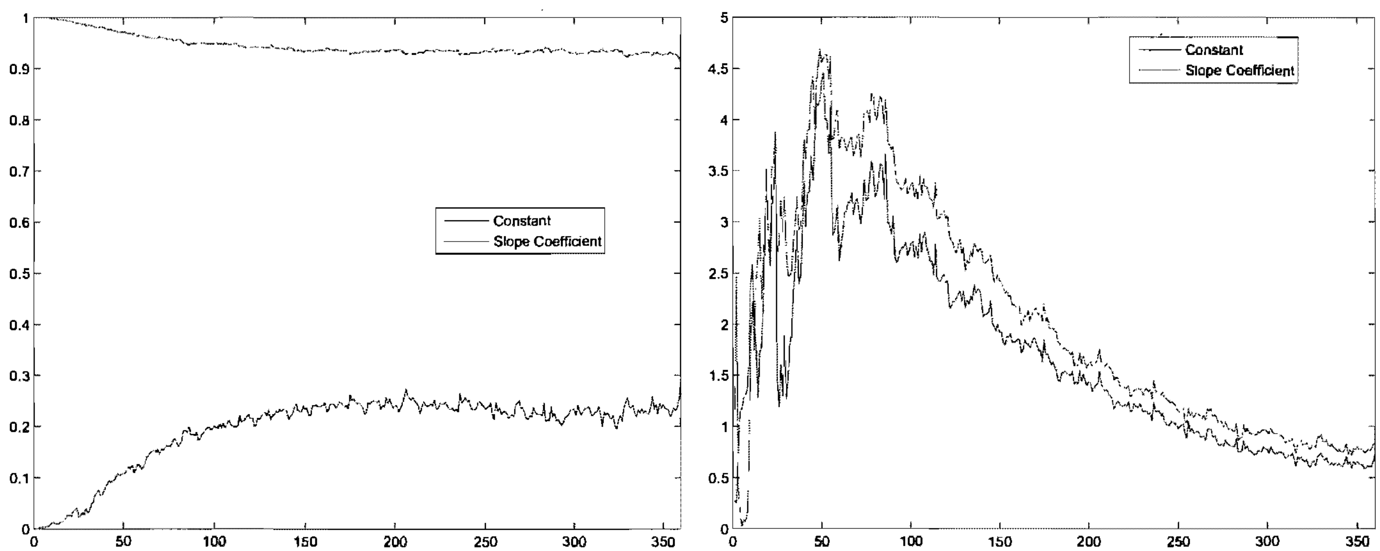
(b) RMSE from each regressions

Figure 2.7: Coefficients from the model-based forecast regressions

The figures displays the coefficients α_n and β_n from regressions of realized monthly average of the Target rate, $\bar{r}_{t,n}$, on the model-forecast,

$$\bar{r}_{t,n} = \alpha_n + \beta_n \Psi_{\bar{r}}(\hat{X}_t, n) + \epsilon_{t,n},$$

where $\Psi_{\bar{r}}(\hat{X}_t, n)$ is the model's forecast and $\hat{X}_t$ are filtered state variables. Panel (a) displays the estimates of α and β across horizons. Panel (b) displays the (absolute value) t-statistics of the null that $\alpha = 0$ and that $\beta = 1$, respectively, for each horizon. In both cases the x-axis is the horizon, h, in the regression, from 1 to 360 days ahead.



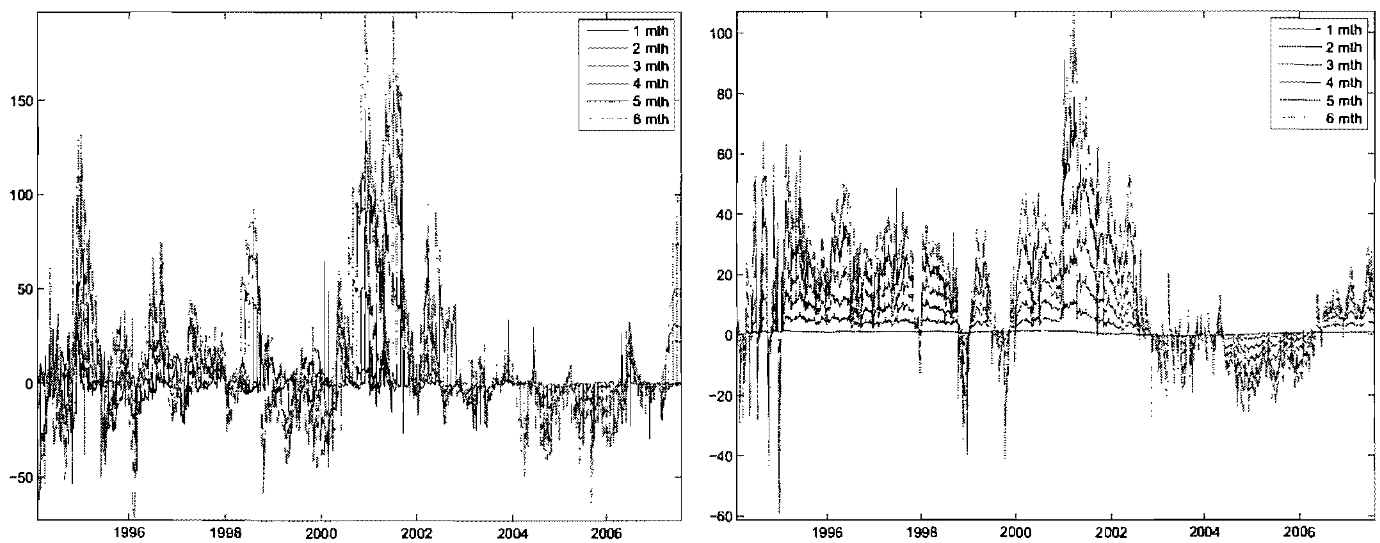
(a) Coefficient estimates

(b) t-statistics

Figure 2.8: Excess returns regressions

These Figures presents actual excess returns, $xr_{t,n}$ on the n-month ahead futures (Panel 2.9a) and predicted excess returns (Panel 2.9b) from regressions of excess returns on the Target rate, the macro factor and the liquidity factor,

$$xr_{t,n} = \alpha + \beta_r r_t + \beta_z z_t + \beta_l l_t + \epsilon_{t,n}.$$



(a) Observed Excess Returns

(b) Predicted Excess Returns

Bibliography

Barolini, L., G. Bertola, and A. Prati (2002). Day-to-day monetary policy and the volatility of the Federal funds interest rate. *Journal of Money, Credit and Banking 34*, 137–159.

Bernanke, B. and K. Kuttner (2005). What explain the stock market's reactions to federal reserve policy. *The Journal of Finance 60*, 1221–1257.

Christiano, L., M. Eichenbaum, and C. Evans (1998). Monetary policy shocks: What have we learned and to what end? WP6400.

Christoffersen, P., K. Jacobs, L. Karoui, and K. Mimouni (2007). Estimating term structure models using swap rates. McGill University.

Cochrane, J. and M. Piazzesi (2002). The fed and interest rates Ů a high-frequency identification. *American Economic Review 92*, 90–95.

Darolles, S., C. Gourieroux, and J. Jasiak (2002). Structural laplace transform and compound autoregressive models. Technical report, CREST.

Das, S. (2002). The surprise element: jumps in interest rates. *Journal of Econometrics 106*, 27–65.

Duffee, G. (2002). Term premia and interest rate forecasts in affine model. *The Journal of Finance 57*, 405–443.

Duffie, D., J. Pan, and K. Singleton (2000). Transform analysis and asset pricing for affine jump-diffusion. *Econometrica 68*, 1343–1376.

Duffie, D. and K. Singleton (1997). An econometric model of the term structure of interest-rate swap yields. *The Journal of Finance 52*, 1287–1321.

Fontaine, J.-S. and R. Garcia (2008). Bond liquidity premia. Working Paper.

Garcia, R. and H. Schaller (1999). Are the effects of monetary policy asymmetric. *Economic Inquiry 40*, 102–119.

Gourieroux, C., A. Monfort, and V. Polimenis (2002). Affine term structure models. Technical report, CREST.

Gurkaynak, R., B. Sack, and E. Swanson (2007). Market based measures of monetary policy expectations. *Journal of Business and Economics Statistics 25*, 201–212.

Hamilton, J. (1996). The daily market for Federal funds. *The Journal of Political Economy 104*, 26–56.

Hamilton, J. (2008). Assessing monetary policy effects using daily federal funds futures contracts. *Federal Reserve Bank of St. Louis Review 90*, 377–393.

Hamilton, J. (2009). Daily changes in Fed funds futures prices. *Journal of Money, Credit and Banking 41*, 567–582.

Hamilton, J. and O. Jordà (2002). A model of the Federal funds rate target. *The Journal of Political Economy 110*, 1136–1167.

Johannes, M. (2004). The statistical and economic role of jumps in coutinous-time interest rate models. *The Journal of Finance 59*, 227–260.

Johnson, N., S. Kotz, and N. Balakrishnan (1997). *Discrete Multivariate Distributions* (First ed.). New York Wiley.

Julier, S. and J. Uhlmann (1996). A general method for approximating nonlinear transformations of probability distributions. Robotics Research Group, University of Oxford.

Julier, S., J. Uhlmann, and H. Durrant-Whyte (1995). A new approach for filtering nonlinear systems. *The Proceedings of the American Control Conference 3*, 1628–1632. Seattle.

Krueger, J. and K. Kuttner (1996). The Fed funds futures rate as a predictor of Federal reserve policy. *The Journal of Futures Markets 16*, 865–879.

Kuttner, K. (2001). Monetary policy surprises and interest rates: Evidence from the Fed funds futures market. *Journal of Monetary Economics 47*, 523–544.

Leippold, M. and L. Wu (2003). Design and estimation of multi-currency quadratic models. *European Finance Review 7*, 47–73.

Piazzesi, M. (2005). Bond yields and the federal reserve. *The Journal of Political Economy 113*, 311–344.

Piazzesi, M. and E. Swanson (2006). Futures prices as risk-adjusted forecasts of monetary policy. Technical Report 2006-23, Federal Reserve Bank of San Francisco.

Polimenis, V. (2006). Skewness corrections for asset pricing. working paper.

Rudebusch, G. (2006). Monetary policy inertia: Fact or fiction? *International Journal of Central Banking 2*, 85–135.

Sack, B. (2004). Extracting the expected path of monetary policy from futures rates. *The Journal of Futures Markets 24*, 733–754.

Skellam, J. (1946). The frequency distribution of the difference between two Poisson variates belonging to different populations. *Journal of the Royal Statistical Society 109*, 296.

Wan, E. and R. V. der Merwe (2001). *Kalman filtering and neural network.*, Chapter The unscented Kalman filter. Wiley and Sons.

Chapitre 3

The Equity Premium and The Volatility Spread

The Equity Premium and The Volatility Spread : The Role of Risk-Neutral Skewness

Bruno Feunou
Duke University

Jean-Sébastien Fontaine
Bank of Canada

Roméo Tedongap
Stockholm School of Economics

Abstract

We introduce the Homoscedastic Gamma [HG] model where the distribution of returns is characterized by its mean, variance and an independent skewness parameter under both measures. The model predicts that the spread between historical and risk-neutral volatilities is a function of the risk premium and of skewness. In fact, the equity premium is twice the ratio of the volatility spread to skewness. We measure skewness from option prices and test these predictions. We find that conditioning on skewness increases the predictive power of the volatility spread and that coefficient estimates accord with theory. In short, the data do not reject the model's implications for the equity premium. We also check the model's implications for option pricing and show that the information content of skewness leads to improved in-sample and out-of-sample pricing performances as well as improved hedging performances. Our results imply that expanding around the Gaussian density is restrictive and does not offer sufficient flexibility to match the skewness and kurtosis implicit in option data. Finally, we document the term structure of option-implied volatility, skewness and kurtosis and find that time-dependence in returns has a greater impact on skewness.

Keywords: Options, Implied Skewness, Risk Premium, Volatility Spread.
JEL Classification: G12, G13.

A previous version of this paper was titled "The Implied Volatility and Skewness Surface". We thank Peter Christoffersen, Redouane Elkamhi, René Garcia, Scott Hendry, Steve Heston, Teodora Paligorova and Jun Yang for comments and discussions. We also thank seminar participants at the Bank of Canada, Financial Econometrics Lunch Group (Duke University), CIRANO 2009 Financial Econometrics Conference, 2009 Econometric Society NASM, EFA 2009 and FMA 2009. *Correspondence:* [redacted]

I Introduction

We propose the Homoscedastic Gamma model [HG] in which innovations of market returns are parameterized by their mean, variance and skewness. The skewness parameter can be chosen independently and we nest the Black-Scholes-Merton [BSM] case if skewness is zero. We follow Christoffersen et al. (2009) and provide a Stochastic Discount Factor [SDF] under which stock returns are HG under both the historical and risk-neutral probability measures. This model delivers a sharp prediction about the relationship between the risk premium, volatility and skewness : the equity premium is equal to twice the *ratio* of the volatility spread to skewness.

The HG model preserves BSM's parsimony and closed-form option prices. Thus, we measure the volatility and skewness implicit in option prices. We can then perform regressions of SP500 excess returns on the ratio of the volatility spread to skewness. We find that coefficients have the correct sign and magnitude, and that conditioning on skewness improves the predictive power of the volatility spread. In short, the data support the model's restrictions on the equity premium. Reversing the relationship, and interpreting the volatility spread as the returns on a portfolio of options, we show that a version of the CAPM conditional on skewness "explains" the returns on the the volatility spread portfolio. This offers an answer to the question posed in Carr and Wu (2009) regarding which factor may explain the variance premium.

An important implication of this new stylized fact is that an understanding of the volatility spread, and its relationship with the compensation for risk, demands an understanding of risk-neutral skewness. Intuitively, both the price of risk and the volatility spread are related to the risk-neutral skewness. The volatility spread has been linked to variance risk (Bakshi and Kapadia (2003), Bollerslev et al. (2008), Carr and Wu (2009)) or to a left-skewed and fat-tailed returns distribution (Bakshi and Madan (2006), Polimenis (2006)).[1] While these different channels explain the volatility spread, they do not have the same implications for risk-neutral skewness. This should help discriminate across competing theories of the observed volatility spread. Clearly, understanding the source of risk-neutral skewness is a key research objective.

As a further check for the importance of risk-neutral skewness, we test

[1] Bakshi and Madan conclude that historical skewness do not play an important role in the determination of the volatility spread but they do not consider risk-neutral skewness.

its pricing implications for option contracts written on the SP500 index. We consider the simple HG model and variants analogous to the practitioner's version of the BSM model [P-BSM and P-HG]. We interpret these variants as expansions around the HG distributions but develop, and impose, restrictions ensuring the identification of the skewness parameter with the true underlying risk-neutral skewness. Overall, HG-based models significantly improve in-sample and out-of-sample performances relative to Gaussian-based models but with the same number parameters or less. They also increase hedging performances at horizons up to 4 weeks.

The results imply that expanding around the Gaussian density is restrictive and does not offer sufficient flexibility to match the skewness implicit in the data. Another way to view the evidence is to consider the results of Bates (2005) and Alexander and Nogueira (2005). Essentially, for any contingent claim that is homogenous of degree one, option partial derivatives with respect to the underlying can be computed, model-free, by taking partial derivatives of option prices with respect to strike prices. In practice, however, a parametric model is fitted to observed prices from which derivatives can be imputed. The relative hedging performances of the P-BSM and of the P-HG models imply that accounting for skewness explicitly offers a better fit of the option price curve across the strike continuum, and a better fit of the true underlying option sensitivities. Still, the improvements come with no increase in implementation costs.

Next, we introduce the implied volatility and skewness *surface*, an extension of the implied volatility curve. Beyond its simplicity and ease of computation, the BSM's implied volatility [IV] curves deliver transparent comparisons of options through time and across strike prices. Repeating the inversion of the IV curve across values of skewness delivers the implied volatility and skewness surface. The surface provides a transparent understanding of IV curve variations in term of skewness. We find that the volatility-skewness relationship is smooth in practice: negative (positive) skewness increases (decreases) the implied volatility of out-of-the-money [OTM] calls and decreases (increases) the implied volatility of in-the-money [ITM] calls. We draw two important conclusions. First, the HG model can restore the symmetry of the observed IV curve. Second, the level of the IV curve also depends on skewness.

Finally, we study the term structure of implied volatility, skewness and excess kurtosis. This is a first step to understand the impact of time dependence on risk-neutral moments. The HG model delivers estimates of risk-neutral volatility and risk-neutral skewness at longer horizons than a non-parametric approach. The evidence suggests that skewness decays at a rate slower than what implied by the i.i.d. assumption. In other words, the time-dependence structure of returns has a larger impact on the term structure of skewness skewness than on volatility and kurtosis. To our knowledge, this differential impact of returns time-dependence on higher moments has never been documented.

Related Literature

The stylized observations that IV curves typically display a smile, a skewed smile or a smirk have been interpreted as evidence of skewness and kurtosis in the underlying risk-neutral distribution of stock price (e.g. Rubinstein and Jackwerth (1998)). In practice, the importance of skewness for pricing stock index options has been highlighted in the empirical works of Bakshi et al. (1997), Bates (2000) and Christoffersen et al. (2006). However, it is generally difficult to invert option prices and obtain estimates of implied volatility or implied skewness. In most cases, volatility and skewness are not independent or, else, option prices are not available in closed-form, rendering inversion computationally expensive. Then, although the increased sophistication allows for a better fit of observed IV curves, our understanding of skewness remains incomplete. In particular, the linkages between skewness, implicit from option prices, the risk premium, measured from equity returns, and the volatility spread remains elusive. The i.i.d. case leads to a stylized model but allows us to maintain parsimony and analytical tractability.

Option pricing based on a Gram-Charlier expansion also offers direct parametrization of skewness and kurtosis (Jarrow and Rudd (1982), Corrado and Su (1996), Potters et al. (1998)). However, approximations of the underlying risk-neutral density often turn negative implying that estimated values of cumulants do not belong to a true distribution. Jondeau and Rockinger (2001) offer a natural remedy and impose a positivity constraint on the estimated density. This is not innocuous. The range of admissible skewness values is restrictive for option pricing applications.[2] Finally, models based on Gram-Charlier do not provide a change of measure linking the historical and risk-neutral measure.[3]

[2] Jondeau and Rockinger (2001) establish that their restriction imply that skewness takes values within $(-1.0493, 1.0493)$. León et al. (2006) establishes the impact of this restriction for option pricing.

[3] Note also that closed-form option prices typically result from a first-order approximation. This may not be relevant in practice for option pricing but the impact of this approximation on estimates of implied skewness has not been discussed.

Bakshi and Madan (2000) provide a non-parametric measure of skewness (and other higher-order moments) implicit from option prices. This was exploited by Bakshi et al. (2003), who focus on measures of skewness in the cross-section and on the link with index skewness. Dennis and Mayhew (2000) consider determinants of the cross-section of skewness and Rompolis and Tzavalis (2008) attribute the bias in volatility regressions to the risk-neutral skewness. Christoffersen et al. (2008) explores the information content of option data for future stock betas. However, the pricing or hedging implications of skewness for option prices cannot be handled within this model-free framework.[4]

The rest of the paper is organized as follow. Section II introduces the Homoscedastic Gamma model [HG] as well as the SDF and contains the main asset pricing implications. In particular, it contains the mapping between parameters under each measure and derives the option pricing function. Section III presents the data. Section IV perform regression-based tests of the model's implications for the equity premium and the volatility spread, and discusses the results in the context of equilibrium model. We introduce a practitioner's analog in Section VI and compare in-sample, out-of-sample and hedging performances of HG and BSM-based models in Section VII. Section V explores the empirical properties of the implied volatility and skewness surface while Section VIII provides estimates of the term structure of volatility, skewness and kurtosis. Section IX concludes.

II The Homoscedastic Gamma Model

This section introduces the Homoscedastic Gamma model for stock returns. The model possesses three crucial properties that makes it a natural choice to study the linkages between the equity premium, the volatility spread and skewness. First, skewness is parameterized directly and is independent of the mean and variance. Second, its density and characteristic functions are known in closed-form. Third, the distribution of returns remains HG for all investment horizons under both the historical and the risk-neutral probability measures whenever the SDF is exponential in aggregate wealth. In particular, this delivers an explicit mapping between moments under each measures. Finally, we obtain closed-form prices for European options of any

[4]Note, also, that this approach requires approximations of integrals over the moneyness domain. Although Dennis and Mayhew (2000) consider theimpact of sampling error under the null of the BSM model, the accuracy of skewness estimates areunknown in the presence of measurement errors or in a non-gaussian setup.

maturity as a function of volatility and skewness. We can then efficiently invert option prices to obtain implied volatility and skewness surfaces. Indeed, when setting skewness to zero our model simplifies to the BSM and we recover the usual BSM implied volatility curve.

A Returns Under the Risk-Neutral Measure

We assume that stock prices, S_t, follow a discrete-time process whereas the logarithm of gross returns, R_t, over an interval of time Δ, say, follows

$$\begin{aligned} R_{t+\Delta} &\equiv \ln\left(S_{t+\Delta}/S_t\right) = \mu^* \; \Delta + \sqrt{\sigma^{*2}\Delta} \; \varepsilon^*_{t+\Delta} \\ \varepsilon^*_{t+\Delta} &\sim SG(\alpha^*\left(\Delta\right)), \end{aligned} \tag{3.1}$$

under the risk-neutral measure where μ^* and σ^{*2} are the risk-neutral drift and variance, respectively. Return innovations, $\varepsilon^*_{t+\Delta}$, follow a Standardized Gamma [SG] distribution with zero mean, unit variance and skewness α^*. The SG distribution is defined in terms of the Gamma distribution, $\Gamma(k,\theta)$, as

$$X \sim SG(\alpha) \Leftrightarrow \frac{2}{\alpha}(X + \frac{2}{\alpha}) \sim \Gamma\left(\frac{4}{\alpha^2}, 1\right), \tag{3.2}$$

where the scale parameter is fixed to $\theta = 1$. Given that the Gamma definition has mean $k\theta$, variance $k\theta^2$ and skewness $2/\sqrt{k}$, it follows that one-period returns in the HG model have mean $\mu^*\Delta$, variance $\sigma^{*2}\Delta$ and skewness $\alpha^*(\Delta)$. We express skewness as function of Δ to reflect the choice of the interval's length. A key simplifying assumption is that the conditional distribution of returns does not vary through time. Still, the model could be thought as holding conditionally, with parameters μ_t, σ_t and α_t indexed by time.

B Returns Under The Historical Measure

We provide a change of measure for which the historical distribution of stock returns also belongs to the HG family. The result holds when the SDF is exponential-affine in aggregate wealth returns, which is the case in economies with power utility. Under this assumption, we obtain transparent interpretations of risk-neutral moments in terms of the historical moments and of the compensation for risk. In the HG case, the risk-neutral volatility is greater than the historical volatility when the equity premium is positive and skewness is negative. Also, the volatility spread increases with the equity premium and with the negative asymmetry of returns. When skewness is zero, and returns are Gaussian, only the mean is shifted and the variance is the same under both measures.

First, assume that aggregate returns follow a HG distribution under the historical measure

$$R_{t+\Delta} \equiv \ln\left(S_{t+\Delta}/S_t\right) = \mu\ \Delta + \sqrt{\sigma^2\Delta}\ \varepsilon_{t+\Delta}, \tag{3.3}$$

where $\varepsilon_{t+\Delta} \sim SDG(\alpha(\Delta))$. Next, define the SDF as

$$M_t = \exp\left(-\nu\left(\Delta\right)\varepsilon_t + \Psi\left(\nu\left(\Delta\right)\right)\right), \tag{3.4}$$

for some ν and where Ψ is the logarithm of the conditional moment generating function of $\sqrt{\sigma^2\Delta}\ \varepsilon_{t+\Delta}$. Then, this SDF defines an Equivalent Martingale Measure (EMM), under which the discounted stock price is a martingale, for a unique ν, as stated in the following proposition.

Proposition 1 *If stock returns follow Equation 3.3 and if the Stochastic Discount Factor belongs to the class defined by Equation 3.4 for some ν, then, this SDF defines an Equivalent Martingale Measure for discounted stock prices if and only if*

$$\nu(\Delta) = -\frac{2}{\alpha\left(\Delta\right)\sqrt{\sigma^2\Delta}} + \frac{g\left(\Delta\right)}{g\left(\Delta\right)-1}, \tag{3.5}$$

where

$$g\left(\Delta\right) = \exp\left(-\frac{(\mu-r)\Delta}{4}\alpha\left(\Delta\right)^2 + \frac{\alpha\left(\Delta\right)\sqrt{\sigma^2\Delta}}{2}\right),.$$

See the Appendix for all proofs. This is a direct application of results from Christoffersen et al. (2009). Note that the price of risk, $\nu(\Delta)$, converges to the usual result, $(\mu-r)/\sigma^2$, when skewness tends to zero. Also, this result does not imply that the EMM is itself unique but that only one solution exists within the class defined by Equation 3.4.

The following Proposition establishes that stock returns are HG under both measures and characterizes the link between parameters of returns dynamics under each measure.

Proposition 2 *If stock returns under the risk-neutral measure follow Equation 3.3 and if the Stochastic Discount Factor is as in Equation 3.4 for ν given in Proposition 1 then stock returns are given by Equation 3.2 and 3.3 under the risk-neutral and the historical measure, respectively, with* $\varepsilon_t^* =$

$\varepsilon_t - E^Q_{t-1}[\varepsilon_t]$ and where parameters under both measures are linked as

$$\begin{aligned}\sigma^*(\Delta) &= \frac{g(\beta(\Delta)) - 1}{\beta(\Delta)g(\beta(\Delta))} \\ \mu^*(\Delta) &= \mu + 2\frac{\sigma^* - \sigma}{\alpha^*(\Delta)\sqrt{\Delta}} \\ \alpha^*(\Delta) &= \alpha(\Delta)\end{aligned}$$

where we use $\beta(\Delta) = \alpha(\Delta)\frac{\sqrt{\Delta}}{2}$ to simplify the notation. Note that we have $\sigma^* \to \sigma$ and $\mu^* \to \mu + \frac{1}{2}\sigma^2$ when $\alpha - \alpha^* \to 0$.

Due to risk-aversion and non-normality in returns, the risk-neutral volatility differs from its historical counterpart at any horizon. The volatility spread depends on the degree of returns asymmetry, $\alpha(\Delta)$ and the degree of risk aversion through the risk-premium, $(\mu - r)$, implicit in $g(\cdot)$. Whenever skewness is negative and the equity premium is positive, the risk-neutral volatility is greater than the historical volatility (i.e. $\sigma^* > \sigma$). These results are consistent with Bakshi and Madan (2006) and Polimenis (2006). Finally, because of the specific choice of SDF, the risk neutral skewness is the same as the historical skewness.[5]

To see the relationship between ν and skewness, consider a first-order expansion of Equation 3.5 around $\alpha(\Delta) = 0$. For small deviations around the symmetric case we have

$$\nu(\Delta) \approx \frac{\mu - r}{\sigma^2} + \frac{1}{2} + \frac{(\mu - r)^2 + \frac{\sigma^4}{12}}{\sigma^3}\beta(\Delta), \tag{3.6}$$

Note that $\nu(\Delta)$ tends toward the usual result, $\frac{\mu - r}{\sigma^2}$, when skewness approaches zero. Then, as expected, ν can be interpreted as the price of risk. Moreover, it is a function of the equity risk premium, of the volatility and of skewness.

Another way to see the link between the equity premium and the volatility spread is to note that

$$\ln(S_{t+\Delta}/S_t) = \mu\Delta + 2\frac{\sigma^* - \sigma}{\alpha^*\sqrt{\Delta}} + \sqrt{\sigma^{*2}\Delta}\varepsilon^*_{t+\Delta},$$

where the middle term converges to zero as skewness approaches zero.[6] Tak-

[5]One can show that an SDF exists such that the returns distribution belongs to the HG family under both measures with *both* the variance and the skewness parameter shifted. However, this SDF is not in general within the exponential-affine class and the link between moments is not transparent.

[6]In the limit, as skewness becomes zero, stock returns follow the usual square-root

ing expectations and re-arranging reveals the following important restriction between the equity premium, the volatility spread and the risk-neutral skewness,

$$E_t^P[\ln(S_{t+\Delta}/S_t)] - E_t^Q[\ln(S_{t+\Delta}/S_t)] = -2\frac{\sigma^* - \sigma}{\alpha^*\sqrt{\Delta}}. \tag{3.7}$$

In the HG model, the volatility spread is solely due to the presence of skewness and not to volatility being priced. Indeed, the volatility spread *and* the equity premium increase when skewness is more negative. In particular, regressions of excess returns on the ratio of the volatility spread to skewness should be more informative than the spread itself. Moreover, the constant is zero and the predicted value for the coefficients is -2. This provides a simple test that we implement below.

C Option Prices

We are now ready to provide a closed-form price for European style contingent claims on a stock. This simple homoscedastic model is stable under temporal aggregation. That is, if returns over two successive intervals follow a SDG distribution then returns over the sum of the intervals also follow a SDG distribution. This is a key property to obtain closed-form option prices for all maturities. Consider (log) stock returns over an arbitrary investment horizon H. Define $M \equiv \frac{H}{\Delta}$ as the number of time steps over this horizon. Then,

$$\begin{aligned} R_{t,M} &\equiv \textstyle\sum_{j=1}^{M} R_{t+j\Delta} = \ln(S_{t+\Delta M}/S_t) \\ &= \mu^* M\ \Delta + \sigma^* \sqrt{\Delta M}\ \varepsilon^*_{t,M}, \end{aligned}$$

where the return innovation, $\varepsilon^*_{t,M}$, is given by[7]

$$\varepsilon^*_{t,M} \equiv \sum_{j=1}^{M} \frac{\varepsilon^*_{t+j\Delta}}{\sqrt{M}} \sim SDG(\alpha^*(\Delta)/\sqrt{M}).$$

A no-arbitrage price, $C_t(K, H)$, of a European call option with strike price K and maturity H can be obtained from the discounted risk-neutral expectation of the terminal payoff,

$$C_t(K,H) = E_t^Q\left[\exp(-rH)\max\left(S_{t+H} - K, 0\right)\right].$$

As usual, the solution is function of the other model parameters: the risk-

process.

[7]This follows directly from the fact that the Gamma distribution is summable.

free rate, r, the risk-neutral volatility, $\sigma^*(\Delta)$, and the scaled skewness $\beta(\Delta)$. Moreover, the solution depends on the direction of asymmetry. Specifically, the case with no skewness corresponds to the BSM formula while we have the following proposition otherwise.

Proposition 3 *If the logarithm of gross stock returns follows a Homoscedastic Gamma process under the risk-neutral measure, as in Equation 3.2, then the price of a European call option is*

$$C_t(K,H) = S_t C_{1,t} - e^{(-rH)} K C_{2,t}, \tag{3.8}$$

where, if the skewness is negative (i.e. $\alpha(\Delta) < 0$),

$$C_{1,t} = P\left(\frac{H}{\beta(\Delta)^2}, d_1(\Delta)\right) \tag{3.9}$$

$$C_{2,t} = P\left(\frac{H}{\beta(\Delta)^2}, d_2(\Delta)\right), \tag{3.10}$$

and, if the skewness is positive, (i.e. $\alpha(\Delta) > 0$),

$$C_{1,t} = Q\left(\frac{H}{\beta(\Delta)^2}, d_1(\Delta)\right) \tag{3.11}$$

$$C_{2,t} = Q\left(\frac{H}{\beta(\Delta)^2}, d_2(\Delta)\right), \tag{3.12}$$

The functions $P(a,z)$ and $Q(a,z)$ are the regularized gamma functions[8] *defined by*

$$P(a,z) = \frac{\gamma(a,z)}{\Gamma(a)}$$

$$Q(a,z) = \frac{\Gamma(a,z)}{\Gamma(a)},$$

respectively, with $\gamma(a,z)$ and $\Gamma(a,z)$ the upper and the lower incomplete gamma functions[9] *and where d_1 and d_2 are defined as*

$$d_2(\Delta) = \frac{\ln(K/S_t) - \left(r_f + \frac{\ln(1-\beta(\Delta)\sigma^*(\Delta))}{\beta(\Delta)^2}\right) H}{\beta(\Delta)\sigma^*(\Delta)},$$

$$d_1(\Delta) = d_2(\Delta)(1 - \beta(\Delta)\sigma^*(\Delta)).$$

[8] We use the standard notation for the regularized gamma functions, $P(a,z)$ and $Q(a,z)$, possibly at the cost of some confusion with the usual notations for the historical and risk-neutral probability measures P and Q.

[9] Note that we have $P(a,z) + Q(a,z) = 1$, which is a convenient property when computing derivatives (see below).

III Data

This section introduces the data and presents some summary statistics. We use prices of call options on the S&P500 index observed on each Wednesday in the period from 1996 to 2004. Using Wednesday observations is common practice in the literature (e.g. Dumas et al. (1998)) to limit the impact of holidays and day-of-the-week effects. Consequently, the return horizon in Equation 3.2 is set to one week in the following. We exclude observations with less than 2 weeks to maturity, no bid available or with zero transaction volume. We also filter observations for violation of upper and lower pricing bounds on call prices.

Next, we introduce a second sample that group option prices at the monthly frequency. This reduces the noise in the estimates of volatility and skewness used in excess returns regressions. Another benefit of this approach is that it ensures enough observations to estimate our model in each maturity group. This allows us to draw the implied volatility and skewness surface in different maturity groups and, as a byproduct, to obtain a term structure of skewness and volatility. To group observations, we use settlement dates rather than calendar months. Since each contract settles on the third Friday of a month, we group all observations intervening between two successive settlement dates.[10] All weekly observations occurring within such a sub-period can be unambiguously attributed to one maturity group.[11] Note that settlement dates follow a regular pattern though time: contracts are available for 3 successive months and then for the next 3 months in the March, June, September, December cycle. This leads to maturity groups with 1, 2 or 3 months remaining to settlement and then between 3 and 6, between 6 and 9, and between 9 and 12 months remaining to settlement.[12]

Table I displays the number of contracts, the average call price and the average implied volatility across moneyness (Panel (a)), across maturity (Panel (b)), and a detailed cross-tabulation across moneyness and maturity (Panel (c)). The Black-Scholes IV curve is asymmetric in the overall sample, displaying a rising pattern with moneynesss, and signaling a sharp left skew in the risk-neutral distribution of returns. Also, the IV curve is flat, or

[10]These subperiods have varying length depending on the (calendar) months they cover.

[11]Take any contract, on any observation date. This contract is assigned to the 1-month maturity group if its settlement date occurs on the following third-Friday, to the 2-month group if it occurs on the next to following third-Friday, etc.

[12]Within a given month, and within a given maturity group, the same contract (i.e. same strike price) is observed with successively shorter maturities. However it is priced consistently under the null of i.i.d. returns innovations throughout the month.

slightly decreasing, with maturity. Disaggregation reveals variations in the shape of the IV curve at different maturities. Starting from the shortest maturity, the IV curve initially follows an asymmetric smile with higher volatility values for in-the-money options. Hereafter, the asymmetry increases as we consider longer maturities and the (average) IV curve eventually becomes monotone in moneyness for the longer maturities.

Note that the composition of the sample varies with maturities. Out-of-the-money contracts dominate for long maturities while in-the-money contracts dominate for short maturities. This is due to the issuance pattern of new option contracts. Newly issued, long-maturity call options are typically deep-out-the-money, in anticipation of the index upward drift through time. As we consider shorter maturities, the composition becomes more balanced. At the shortest horizon, most call options are deep in-the-money, since the exchange does not regularly issue short horizon out-of-the-money call options. This implies that the average IV curve reflects, in part, a composition bias with most in-the-money options having short maturities and most out-of-the-money options having long maturities. Because short maturity options have higher implied volatility on average, this makes the average IV curve more smirked.[13] Finally, Panel (a) of Figure 3.1 presents the number of available observations for each day, which averages around 40 and typically ranges between 20 and 50 contracts. Panel (b) decomposes this number and presents the proportion of contracts in each moneyness category.

IV The Volatility Spread And The Equity Premium

A Model's Implications

When the representative SDF can be approximated by the exponential-shift given in Equation 3.4 we have a tight link between the price of risk, the volatility spread and skewness. After some manipulation of Equation 3.7, we obtain

$$\ln\left(S_{t+i}/S_t\right) - r^{(i)} - \omega_t^{(i)} = -2\frac{\sigma^{*(i)} - \sigma^{(i)}}{\alpha_t^{(i)}} + \sigma^{*(i)}\varepsilon_{t+i}^{*},$$

[13]This highlights the importance of using a model that can handle maturity differences. In particular, models based on density approximation are not robust to this composition effect.

for an investment horizon i and where $r^{(i)}$ is the risk-free rate for that horizon and ω_t is the Jensen adjustment term.[14] In the following, we test this implication of the HG model and its ability to capture the volatility spread and the equity premium. We perform regressions of SP500 (log) excess returns on the ratio of the volatility spread to skewness. The key predictions are that the constant should be zero and that the coefficient should be -2.

B Aggregating Data

We obtain estimates of risk-neutral volatility and skewness from option data. Estimates of skewness for different maturities are noisy in weekly data. This is in part due to the number of option prices available each week in each category. One simple solution is to group price observations at the monthly level where we define a month as the period between successive expiration dates which occur every third Friday (See Section III). Within each month, we have repeated observations of the same contracts over a period of 4 (or 5) weeks.[15] This implicitly assumes i.i.d. return innovations throughout a month, which is consistent with the model and reasonable over this short time span. It also implies that the maturity *date* of each contract is constant throughout each month and, thus, that the skewness estimate pertains to a set of contracts that mature at fixed maturities. Finally, we measure the historical volatility using the observed realized volatility.

We estimate our preferred version of the model each month through minimization of squared pricing errors.[16] Figure 3.2 presents the time series of our volatility estimates (Panel (a)) and of our skewness estimates (Panel (b)). Skewness typically varies around -1 but dipped close to -2.5 in the summer of 1998 and in the second half of 1999, and slightly below 1.5 in the Fall of 1996 and the Spring of 2004.

C Implied Skewness And The Risk Premia

Table II presents the results from regressions of excess returns at horizons of 1, 3, 6, 12, 24 and 36 months on the ratio of the volatility spread to

[14]This term is a function of both skewness and volatility but the first term of its Taylor expansion is the usual correction in the Gaussian case, $\frac{1}{2}\sigma^2$.

[15]Some contracts are not observed each Wednesday within a month. New contracts become available to participants as the index moves away from the range of available strike prices. Also, some contracts are not available each week because they were excluded from the weekly sample due to liquidity concerns.

[16]Specifically, we estimate a restricted version of the practitioner's HG model that allow for kurtosis but maintain the identification of the risk-neutral volatility and skewness (See Section VII). As a robustness check (not reported) we repeated the exercise using skewness estimated from the simple HG model presented above. The results are not qualitatively different.

skewness. [17] The results are striking. Point estimates for the slope coefficient are close to -2 as predicted by the model. Moreover, at horizons of 3, 6, and 12 months, where we would expect the forward-looking nature of the option-implied estimate to be the most relevant, estimates are -2.24, -2.04 and -2.13, respectively. In fact, at any horizon, we cannot reject the null hypothesis that the coefficient is equal to -2. Next, the constant is not significantly different from zero so that the two most important implications of the model cannot be rejected empirically. Finally, the predictability of excess returns is low at the 1-month horizon (i.e. R^2 is 1.85%) but rises steadily with the horizon, reaching 5.6%, 9.7% and 11.3% at horizons of 6, 12 and 36 months.

For comparison with results available in the existing literature, we also consider regressions on the volatility spread which displays some predictive power at horizons of 9 and 12 months. However, coefficients are not significant at other horizons. Finally, we ask if the volatility spread contains information beyond that revealed by the volatility to skewness ratio. The results from the regressions are presented in Table II. Since volatility and the ratio of the volatility spread to skewness are correlated, the coefficients become unreliable, even changing signs. However, their combined predictive power does not rise above that of the volatility to skewness ratio, further supporting the implications of the model.

D Discussion

We can also interpret the results in the broader context of a general equilibrium model. There, the price of risk is determined by preference parameters. In particular, in an economy with power utility, ν corresponds to the risk-aversion parameter (see e.g. Bakshi et al. (2003)) which can be estimated given estimates of the risk premium, $\mu - r$, and return volatility, σ, obtained from observed returns data. Equation 3.6, which is repeated here,

$$\nu + \frac{1}{2} \approx \frac{\mu - r}{\sigma^2} + \frac{1}{2}\frac{(\mu - r)^2 + \frac{\sigma^4}{12}}{\sigma^3}\alpha^*,$$

shows that ignoring skewness (the last term) leads to upward bias in the estimate of the price of risk and, hence, of risk aversion. Intuitively, when agents are risk-averse, and the risk premium is positive, a more negative value of skewness corresponds to an increase in the quantity of risk: the probability of lower returns increases. Then accounting for skewness reduces

[17] Precisely, our measures of risk-neutral moments pertain only to the distribution of returns at a horizons of 12 months or less. Nonetheless, if these moments exhibit persistence, their predictive power will extend to longer horizons as is indeed the case

the price of risk required to fit the observed equity premium and, ultimately, leads to lower estimates of risk aversion in the economy.

Note that the effect of skewness is economically significant. Since 1980, the sample mean and volatility of one-year returns is 14.72% and 6.13%, respectively, and the first term of Equation 3.6 is equal to 20.5. In other words, if risk is summarized by the volatility of market returns, then the equity premium appears too large and leads to excessively high estimates of the coefficient of risk aversion. However, the coefficient of skewness, α, in the last term is 12.88. For a value of skewness, say, of -1, the estimate of the price of risk is 7.63, less than half than if we ignore the impact of skewness. Moreover, the estimates of skewness we obtain below are often lower than -1.

The results shows the linkages implied by the HG model between the equity premium, the volatility spread and the skewness hold (Equation 3.7). This suggests that an understanding of the volatility spread and of the equity premium demands an understanding of the determinants of skewness. Moreover, it shows that properly conditioning on implied skewness is key to deciphering the information content of options prices for future returns. In fact, reversing the relationship, and interpreting the volatility spread as the returns on a specific portfolio of options,

$$\sqrt{h_t^*} - \sqrt{h_t} = \frac{1}{2\alpha_t^*}\left(E_t^P\left[\ln\left(S_{t+i}/S_t\right)\right] - E_t^Q[\ln\left(S_{t+i}/S_t\right)]\right)$$

we see that a version of the CAPM conditional on skewness "explains" the returns on the volatility spread portfolio. This offers an answer to the question posed in Carr and Wu (2009) which asks what factor may explain the volatility spread.

Our results contrast with existing results (e.g. Bakshi and Kapadia (2003), Bollerslev et al. (2008)) where the spread is linked to variance risk being priced. In our model, the asymmetry in returns shifts the risk premium *and* the risk-neutral volatility. This induces the link between the volatility spread and the equity premium. Similarly, Polimenis (2006) and Bakshi and Madan (2006) link the volatility spread to higher order moments of the *historical* distribution. From the tight linkages we uncover, we conclude that an understanding of the volatility spread, and its relationship with the compensation for risk, demands an understanding of skewness variations. In particular, this new stylized fact should help discriminate across competing theories of the observed volatility spread.[18]

V Implied Volatility and Skewness Surface

In the context of the BSM model, it was recognized early that inverting option prices for the volatility parameter provided a good measure of future returns volatility. However, the HG model offers a separate parametrization for volatility and skewness allowing us to easily measure both the volatility and skewness implicit in option prices.[19] In this section, we study the trade-offs involved between volatility and skewness when fitting option prices. We first analyze how the implied volatility curve varies across different values of skewness and, second, how the implied skewness curve varies with volatility. The results are intuitive. The impact of skewness on implied volatility is asymmetric, depending both on the sign of skewness and of moneyness. In particular, negative skewness tilt a smirked IV curve toward a symmetric smile. On the other hand, the impact of volatility on implied skewness displays a more complex pattern.

An important conclusion from this section is that the HG model exhibits enough flexibility to restore the symmetry of the volatility smile. In other words, variations of the IV curve can be interpreted directly in term of skewness within the HG model. Moreover, both the level and the shape of the IV curve are sensitive to the choice of the skewness parameter. In particular, this implies that empirical studies of the volatility spread based on BSM implied volatility are affected by measurement errors due to the impact of skewness.

A Inverting The Implied Volatility and Skewness Surface

Volatility and skewness cannot be inverted uniquely from a single option price. Instead, for each strike price, the HG model implies a function describing the set of volatility and skewness pairs matching the observed price: a volatility-skewness curve. This is in contrast with the BSM model where any given option price can be inverted uniquely for the volatility parameter. Of course, if the HG model is true, using options with different strike prices would identify uniquely a volatility-skewness couple. In fact, only two dif-

[18] Bakshi and Madan (2006) conclude that the historical skewness plays a relatively small role in the determination of volatility spread but they did not consider risk-neutral skewness.

[19] See Bates (1995) for a review of the literature on the forecasting of volatility using option prices and Andersen et al. (2005) for a review of volatility measurement from stock returns. See Kim and White (2003) for a discussion of the lack of robustness of the usual sample skewness estimator

ferent strike prices would be sufficient for this purpose. In practice, the HG model extends the BSM model in only one direction, allowing for a skewness parameter. Other deviations from the underlying assumptions cause the volatility-skewness curve to vary across moneyness in such a way that no unique couple can match every observed price. Thus, in the HG model, the counterpart to the IV curve is the implied volatility and skewness *surface.* This surface is the representation of the set of volatility and skewness pairs matching the observed option prices for varying strike prices.

To draw the volatility and skewness surface, we first pick a value of skewness from a grid. Then, each day and for each available strike price, we invert the option price for the volatility parameter and obtain an implied volatility curve. As we vary the value of skewness we obtain different IV curves and, together, they yield an implied volatility and skewness surface. A section of this surface at a given value of skewness is one possible IV curve. Each day, these different IV curves are alternative, and equivalent, representations of the data. Each embodies all the information about the distribution of returns and, in addition, measurement errors due to transaction costs, illiquidity and asynchronous trading. The next section provides the results.

B Impact Of Skewness on Implied Volatility Curves

The average volatility-skewness surface is given in Figure 3.3 in level (Panel (a)) and in percentage deviations from the benchmark BSM IV curve (Panel (b)). Panel (a) displays the usual smirk in the IV curve when skewness is zero. More interestingly, it shows that the average IV curve is flat for values of skewness around -1.[20] Next, consider the deviations from the BSM curve in Panel (b). The case with skewness equal to zero corresponds to a straight line at zero. As we consider values of skewness away from zero, the IV curve is tilted one way or another depending on the sign of return asymmetry considered. For negative values of skewness, the IV curve is tilted toward positive values of moneyness. Conversely, for positive values of skewness, the IV curve is tilted toward negative values of moneyness. In other words, as we shift probability mass toward the left (right) tail of the return distribution, the implied volatilities required to match observed prices increase (decrease) for out-of-the-money calls and decreases (increases) for in-the-money calls thereby tilting the IV curve back toward a symmetric smile. In the extreme cases, allowing for non-zero skewness can raise or decrease measured implied volatility by more than 15% relative to the BSM case. Clearly, the HG model

[20]The curve is not strictly flat and this may be due to the impact of kurtosis, or to a composition effect. We discuss these possibilities below.

is sufficiently flexible to capture the skewness implicit in option prices.

C Results For Different Option Maturities

Next, Figures 3.4 (a)-(e) present implied volatility and skewness surfaces within different maturity groups while Figures 3.5 (a)-(e) report the same results but in percentage deviations from BSM values. Starting with skewness equal to zero, which corresponds to the BSM case, we see the the shape of IV curve varies substantially across maturities. As discussed in section III, the average BSM IV curve is a slightly asymmetric smile for short maturities: implied volatility obtained from in-the-money options is higher than for out-of-the-money options. The smile then gradually disappears as we increase maturity and the IV curve eventually becomes smirked. For negative values of skewness, and for any maturity, the IV curve is tilted toward a symmetric smile. For short maturities, small negative values of skewness ar sufficient to establish a symmetric smile. As we increase maturity, however, more negative values are necessary. Looking at deviations from the case with zero skewness (Figure 3.5) we see that the impact of a given variation in skewness decreases as we increase maturity.

D Impact Of Volatility On Implied Skewness

Figures 3.6 (a)-(f) present implied values for skewness across different values of implied volatility. For at-the-money options, there is no tradeoff between volatility and skewness. However, the impact of volatility on implied skewness is asymmetric and highly nonlinear on both sides of the moneyness spectrum. As the volatility of returns decreases, and the probability mass is closer to the mean, the skewness value required to match observed price increases for out-of-the-money options, implying a higher right-tail, but decreases for in-the-money options, implying a lower left-tail. The reverse is true when we increase the value of volatility. In both cases the impact is not monotonic as we move away from at-the-money. Rather, the pattern follows a sharp V-shape, or inverted V-shape, where changes of volatility have no impact on implied skewness for at-the-money options, the largest impact for intermediate moneyness, and a lower impact for distant moneyness. This is likely an indication of a trade-off between the skewness and the kurtosis in the HG distribution to match observed prices. Finally, the impact of volatility on implied skewness rises with the option maturity.

VI Practitioner's Models

The previous section shows that the implied volatility and skewness surface can be described as the smooth tilting of the IV curve across values of skewness. However, while the HG model provides enough flexibility to match the skewness present in option data, the IV curve typically remains slightly curved. This is may due to excess kurtosis[21] and may bias our estimates of skewness. In this section, we propose HG-based option pricing formula that are robust to the presence of excess kurtosis. Intuitively, we consider one-term expansions of the HG model that allow for kurtosis. This results is the analog of rationalizations of the P-BSM model as a two-term expansion around the Gaussian density.

The practitioner's variants of the BSM model [P-BSM] and of the HG model [P-HG] capture deviations from the Gaussian or HG distributions by modeling volatility as a quadratic function of moneyness. That is, in the P-BSM case, we have

$$\sigma(\xi) = \sigma_0(\alpha, \kappa)(1 + \gamma_1(\alpha, \kappa)\xi + \gamma_2(\alpha, \kappa)\xi^2),$$

and, in the P-HG case, we have

$$\sigma(\xi) = \sigma_0(\kappa)(1 + \gamma_1(\kappa)\xi + \gamma_2(\kappa)\xi^2)$$

where ξ is moneyness and α and κ are the skewness and excess kurtosis of the risk-neutral distribution, respectively.

The practitioner's IV curve smooths through the cross-section of option prices, ignores local idiosyncracies and focuses on the impact of higher-order moments. This approach is pervasive because of its empirical performance and, also, because its parameters (i.e. σ_0, γ_1 and γ_2) are usually interpreted in terms of the variance, skewness and kurtosis of the true underlying risk-neutral distribution. For these reasons, parameters of the IV curve are commonly estimated without restrictions. In the following, we document that estimates of σ_0, γ_1, and γ_2 vary when we allow for skewness. This contrasts with the usual interpretation of γ_1 as a measure of skewness. The remainder of the section provides restrictions on parameters of the IV function such that we can recover direct estimates of α and κ from option prices.

[21] In contrast with the Gaussian case, the kurtosis of the HG distribution varies with parameter values. Its kurtosis is proportional to the square of the skewness.

A Unconstrained IV Curves

We evaluate empirically the impact of skewness on estimated IV curves. To do so, fix the value of α and estimate the P-HG model at each date. That is, choose values of σ_0, γ_1 and γ_2 that minimized squared pricing errors. Next, average the unconstrained estimates through time. Finally, repeat the exercise for different values of skewness and trace the relationships between skewness and estimates of σ_0, γ_1 and γ_2 . For simplicity we define

$$\xi = \frac{\ln(S/K)(-r\tau)}{\bar{\sigma}\sqrt{\tau}},$$

and group maturities.

Figure 3.7 presents the results. Panel (a) presents average estimates of σ_0. For contracts maturing at the next settlement date, at-the-money implied volatility is 20% on average when skewness is zero. When skewness decreases to -3, estimates of at-the-value volatility increase to 23%. Intuitively, shifting some probability mass toward one side of the distribution involves a trade-off for pricing in-the-money versus out-the-money options. For a constant level of skewness, this tension can be reduced by an increase in the level of volatility. A similar pattern occurs at longer maturities, but the impact of skewness gradually decreases. Panel (b) presents the results for the asymmetry parameter. In line with intuition we find that $\hat{\gamma}_1$ varies linearly with the value of β : both parameters are measures of the underlying skewness. Finally, Panel 3.8c shows that $\hat{\gamma}_2$ also varies substantially with skewness but the relationship is not linear.[22]

The impact of skewness on the IV curve parameter implies that the information on the underlying risk-neutral moments will be shared across unrestricted parameters estimates. Furthermore, the fact that estimates of α and of γ_1 are (linearly) correlated suggests that they are poorly identified. The following section introduces a framework which lead to restrictions on σ_0, γ_1 and γ_2 such that only $\hat{\alpha}$ can capture the risk-neutral skewness. Absent these restrictions, parameters of the IV curve capture some of the asymmetry in the underlying distribution leading to biased estimates of α. The unambigu-

[22]This contrasts with the theoretical results of Zhang and Xiang (2005). They argue that in the Gaussian case and up to a first-order approximation $\sigma_0(\beta, \kappa)$ is linear in the risk-neutral volatility, $\gamma_1(\beta, \kappa)$ is linear in skewness, and $\gamma_2(\beta, \kappa)$ is linear in kurtosis. However, they assume that the skewness and excess kurtosis of the underlying distribution can be chosen independently while in fact there is a tight link between the two for any given correctly specified density. Moreover, they linearize around the case where $\sigma = 0$ and this may lead to a poor approximation.

ous identification of skewness is necessary to provide a measure of the risk premia from implied volatility and skewness and to evaluate the impact of skewness on option prices.[23]

B HG Model With Excess Kurtosis

We now provide a rigorous justification of the P-HG model when the true distribution displays excess kurtosis. We can characterize sufficient restrictions on the parameters of the IV curve such that $\hat{\beta}$ is identified as the risk-neutral skewness in this more general model as well. In this context, parameters of the IV curve are restricted to (known) functions of excess kurtosis. In other words, any deviation from a flat IV curve can only be linked to deviations of κ from zero. As a by-product, we obtain an estimator of the kurtosis in excess of the Gamma distribution.

Intuitively, we assume that the true density of returns can be represented by an Edgeworth expansion around the Gamma distribution. This is similar to earlier work using the Gaussian distribution (Jarrow and Rudd (1982), Corrado and Su (1996)) but the Gamma distribution allows an exact match of the first three moments. We then impose the equality of the option pricing formula under the true model and the P-HG model for at-the-money options.

Suppose that the true evolution of stock returns under the risk neutral measure can be described as

$$R_T = (r - \delta^*)\, T + \sigma^* \sqrt{T} y,$$

where δ^* is a risk-adjustment term, y is a random variable with mean zero, unit variance, skewness, α^* and kurtosis, λ^*. We allow for non-normality beyond the HG and assume that the probability density of y is given by

$$f(y) = h(y) + \frac{\lambda_2^* - \frac{3\alpha^{*2}}{\sqrt{T}}}{4!} \frac{d^4 h(y)}{dy^4}, \tag{3.13}$$

where $h(y)$ is the standardized gamma density. This is a one-term Edgeworth expansion of standardized gamma distribution around the case with no excess kurtosis. If y is normally distributed, then $\alpha = 0$ and $\delta = \frac{\sigma^{*2}}{2}$.This approach captures fat tails in excess of the Gamma distribution but ignores deviations beyond the fourth moment. Our objective here is to allow for a non-trivial implied volatility and skewness surface due to excess kurtosis and to derive explicitly the function $\sigma_0(\kappa)$, $\gamma_1(\kappa)$ and $\gamma_2(\kappa)$. Proposition 4 builds on a

[23]Note that merely imposing $\gamma_1(\alpha, \kappa) = 0$ does not identify an estimator of α with skewness.

no-arbitrage argument and provides a closed-form characterization of option prices and of the risk-adjustment term.

Proposition 4 *If the logarithm of gross stock returns has the density given by Equation 3.13, then the price of a call option, $C^*(K,T)$, with maturity T, underlying price S_0 and strike price K is*

$$\begin{aligned} C^*(K,T) &= S_0 P(a^*, d_1^*) - e^{-rT}\left(1 + T^2\sigma^{*4}\kappa_4\right) K P(a^*, d_2^*) \\ &+ \kappa e^{-rT} K \frac{T^2\sigma^*}{\beta^{*3}} \left[-h''(d_2^*) + \sigma^*\beta^* h'(d_2^*) - \sigma^{*2}\beta^{*2} h(d_2^*)\right], \end{aligned}$$

when $\beta < 0$ and

$$\begin{aligned} C^*(K,T) &= S_0 Q(a^*, d_1^*) - e^{-rT}\left(1 + T^2\sigma^{*4}\kappa_4\right) K Q(a^*, d_2^*) \\ &- \kappa e^{-rT} K \frac{T^2\sigma^*}{\beta^{*3}} \left[-h''(d_2^*) + \sigma^*\beta^* h'(d_2^*) - \sigma^{*2}\beta^{*2} h(d_2^*)\right], \end{aligned}$$

when $\beta > 0$. We define the excess kurtosis, $\kappa = \frac{\lambda_2 - \frac{6\beta^{*2}}{T}}{4!}$, and

$$\begin{aligned} d_2^* &= \frac{\ln(K/S_0) - \left[r + \frac{\ln(1-\sigma\beta)}{\beta^2}\right] T + \ln(1 + T^2\sigma^4\kappa)}{\sigma\beta} \\ d_1^* &= d_2(1 - \sigma\beta) \\ a^* &= \frac{T}{\beta^2}, \end{aligned}$$

where h is the density of the standard gamma distribution.

C Identified practitioner's HG

We are now looking for the restrictions on the parameters of the P-HG model such that estimation of β delivers a convergent estimate of the risk-neutral skewness β^*. Zhang and Xiang (2005) provide the restriction for the case where the Gaussian density is used in the approximation. To find the link between the parameters of the P-HG model with parameters of the true distribution, we impose the following restrictions

$$\begin{aligned} C^*(K,T) &= C(K,T) \\ \frac{\partial C^*(K,T)}{\partial K} &= \frac{\partial C(K,T)}{\partial K} \\ \frac{\partial^2 C^*(K,T)}{\partial K^2} &= \frac{\partial^2 C(K,T)}{\partial K^2} \end{aligned}$$

when evaluated at-the-money (i.e. $K = S_0 e^{rT}$). These restrictions are given in the appendix but note that they are trivially satisfied whenever $\kappa = 0$ since in this case the HG model is true and the IV curve is flat for some value of skewness. Of course this corresponds to the case $\sigma_0 = \sigma$, $\beta = \beta^*$ and $\gamma_1 = \gamma_2 = 0$. We linearize the restrictions around this point (i.e. $\kappa = 0$) and obtain

$$\frac{\sigma_0 - \sigma}{\sigma} = A_1(\sigma, \alpha)\kappa \tag{3.14}$$

$$\gamma_1 = B_1(\sigma, \alpha)\frac{\sigma_0 - \sigma}{\sigma} + B_2(\sigma, \alpha)\kappa \tag{3.15}$$

$$\gamma_2 = C_1(\sigma, \alpha)\frac{\sigma_0 - \sigma}{\sigma} + C_2(\sigma, \alpha)\gamma_1 + C_3(\sigma, \alpha)\kappa, \tag{3.16}$$

where the coefficients are given in the appendix.[24] Then, small deviations of the underlying density from a HG distribution lead to deviations from a flat implied volatility and skewness surface. This highlights the impact of excess kurtosis on the estimates of σ_0, γ_1 and γ_2. It also makes clear that deviations from a flat IV curve are only due to excess kurtosis. More importantly, these restrictions ensure that α corresponds to the risk-neutral skewness and that the practitioner's HG model conforms to the true returns density.

VII Option Pricing Results

In this section, we estimate each model and compare their performance. The results show that the HG framework substantially improves in-sample, hedging and out-of-sample performances. The improvements are robust if we impose identification of the skewness parameters, as discussed in the previous section. Indeed, the improvements remain when the only deviation from the simple HG model is a constant adjustment to kurtosis. Out-of-sample, imposing the identifying restrictions does not degrade pricing performance. In other words, a fixed implied volatility and skewness surface combined with variations in skewness delivers most of the in-sample and out-of-sample improvements. We also compare the hedging performance of each model to highlight the importance of skewness. Again, allowing for varying skewness but fixing kurtosis provides significant improvement.

Overall, our approach delivers a reliable measure of skewness while offering improved forecasting and hedging performance. In contrast, the P-BSM

[24] We differ from Zhang and Xiang (2005) who linearize the restrictions around $\sigma = 0$. Arguably, linearizing around the HG distribution is likely to provide a better approximation than linearizing around the deterministic case.

model does not allow for sufficient flexibility to match the skewness implicit in the data and offers lower hedging and out-of-sample performance. While the more general models we consider perform better in-sample, these improvements disappear out-of-sample. This implies that skewness captures most of the persistent deviations from the Gaussian case and that excess kurtosis and other deviations are transitory.

A Description Of Models

We evaluate the basic HG model and the usual P-BSM model. We also include three different versions of the P-HG model based on the quadratic IV curve,

$$\sigma_t(\xi) = \sigma_0(1 + \gamma_1\xi + \gamma_2\xi^2).$$

where the first version, P-HG1, imposes the simple restriction that $\gamma_1 = 0$. This is another way to see that the usual interpretation of γ_1 as a measure of skewness, while intuitive, is misleading. The second model, P-HG2, imposes the restrictions derived in the previous section and delivers an estimate of skewness robust to excess kurtosis. Finally, P-HG3 is unrestricted. This is a simple way to evaluate the cost, in terms of fit, of estimating skewness directly.

We also introduce "smoothed" versions of these models where some parameters of the IV curves are held constant through the sample. First, the smoothed version of the P-HG1 model, labeled SP-HG1, still imposes that γ_1 is zero but holds γ_2 constant through time. Next, SP-HG2 still allows for a flexible fit of skewness through time but keep excess kurtosis constant through time. We include this model as a simple way to evaluate the relative importance of skewness and kurtosis for option pricing and hedging. Finally, the SP-HG3 model imposes the following structure on the IV curve,

$$\sigma(\xi) = \sigma_0(1 + (\gamma_{10} + \gamma_{11}\alpha)\xi + (\gamma_{20} + \gamma_{21}\alpha)\xi^2).$$

which is a simple attempt to implement the observation made in Section VI that parameters of the IV curve vary with skewness. Finally, estimation is performed through minimization of squared pricing errors in the weekly sample.

B In-Sample RMSE

HG And BSM Models

Table III presents in-sample Root Mean Squared Errors [RMSE] where each results is expressed as a percentage of the BSM's RMSE. Panel (a) presents results across moneyness while Panel (b) presents results across maturities. Although the most flexible (i.e. P-HG3) model achieves an RMSE which is 14% of the benchmark, most of the improvement comes from using the HG distribution: the simpler HG model's RMSE is 37% of the BMS's RMSE but with only more parameter measuring skewness.

Practitioner's Variants

Interestingly, even with one extra parameter, the P-BSM does not offer much improvement (35% vs 37%) over the straightforward HG model. The models offer similar results across maturities but their performances differ across strike prices. The P-BSM improves pricing for in-the-money options at the expense of larger errors for other moneyness groups. On the other hand, the P-HG1 and the P-HG2 models achieve RMSEs that are 28% and 23%, respectively, but with the same number of parameters as the P-BSM model. However, in contrast with the P-BSM model, the lower errors for out-of-the-money options are not compensated by higher errors for options that are nearer the money. Thus, models based on the HG distribution appear to offer more flexibility than the practitioner's BSM in choosing risk-neutral skewness and kurtosis but with equal or less parameters.

Although the naive $\gamma_1 = 0$ restriction seems reasonable, it fails in practice with larger RMSE. Comparing models, we see that imposing the correct identification constraints (P-HG2) provides substantial improvement over the P-HG1, especially for short maturity, out-of-the-money call options. Finally, with one more parameter, the P-H3 offers much lower in-sample RSME (14%) than any other model across all moneyness and maturity categories.

Smoothed Coefficients

Smoothed models have less parameters but the SP-HG2 model still improves (31%) over the P-BSM model but with less parameters. This model has the flexibility to fix skewness from date to date but imposes a constant excess kurtosis. That is, deviations of the IV curve from the HG case are kept constant. Thus, in-sample, a flexible fit of the underlying risk-neutral skewness is key while variations in kurtosis are less important. Finally, while more flexible HG-based models improve the in-sample fit, the next section show that this result is not robust out-of-sample, indicating a relatively minor role for information beyond the third moment.

C Out-of-sample RMSE

The improved performance of models based on the HG distribution may be due to over-fitting and may not hold out-of-sample. This section compares the out-of-sample performance of each model. First, we estimate each model from options in a given week.[25] We then fix these parameters and price options observed in the following week. Table IV presents one-week out-of-sample RMSE for each model across strike prices (Panel (a)) and across maturities (Panel (b)).

Out-of-sample, the improvement in fit relative to the BSM decreases for all models. This indicates that some of the deviations from the Gaussian case are transitory. The lowest relative RMSE is now 57%, obtained for the P-HG3 model, with 4 parameters. On the other hand, the worst result is 68%, obtained for the P-BSM model, with 3 parameters. This add to the evidence that the practitioner's version of the BSM model does not properly fit the persistent skewness and kurtosis present in the data. Strikingly, the SP-HG2 model, which uses 2 parameters and fixes excess kurtosis through the sample, actually improves out-of-sample RMSE (64%) over the more flexible P-HG2 and P-BSM models. Some of the variations in excess kurtosis required to match (in-sample) option prices in this category are transitory, degrading out-of-sample performances. Restricting parameters of the IV curve to capture that part of its variations due to skewness improves the out-of-sample fit.

D Hedging Errors

Hedging errors implied by each model may convey more economic significance to risk-managers. Below, we verify that allowing for skewness significantly alter hedging strategy theoretically, and improves hedging results empirically. Also, we verify that any improved hedging performance persists at horizons beyond one week. The SP-HG2 model, with 2 parameters and where skewness is separately identified, offers the next to best performance. This highlights, again, the value of theoretically sound restrictions. Again, we find that the unrestricted P-HG3 model performs best.

Comparing The Greeks

As in the BSM model, we can compute explicitly the sensitivity of option prices to changes in the underlying parameters, including the sensitivity to changes in skewness. We provide these in the appendix. These derivatives depend on the direction of asymmetry and everywhere the symmetric case (i.e. $\beta = 0$) leads to the standard results from BSM. To see the impact of skewness, we draw options sensitivities across strike prices for different values

[25]For smoothed model, we estimate parameters that are held constant through the sample in a first pass.

of skewness. In the computations, we use the average values of volatility, of the interest rate and of the index level. Figure 3.8 presents results for the first and second derivatives with respect to the underlying, Delta and Gamma, as well as the derivative with respect to volatility, Vega. The results are reported in levels in the top panels (Panel (a) to (c)) and in percentage deviations from the symmetric case in the bottom panels (Panel (d) to (f)).

First, the pattern of Delta across moneyness is familiar. The sensitivity is small for deep out-of-the-money options but grows to close to one for deep in-the-money options. Varying skewness does not alter this picture but looking at levels hides significant deviations. At skewness equal to -2.5, which occurs in our sample, short positions in the stock are as much as 20% higher for some out-of-the money or near to at-the money options. Next, the impact on Gamma is dramatic. In the symmetric case, Gamma appears quadratic in moneyness with highest values for at-the-money options. Decreasing skewness lowers Gamma for in-the-money options but increases Gamma for out-of-the-money options. When skewness is -2.5, Gamma is as much as 50% lower then when skewness is zero for in-the-money options and 50% higher for out-of-the-money options. Finally, skewness has an asymmetric impact on the sensitivity of options to variations in volatility. When skewness is -2.5, Vega decreases by more than 20% for out-of-the-money options but increases by nearly 20% for in-the-money options. Clearly, ignoring the impact of skewness can lead to large hedging errors, which is confirmed empirically in the next section.

Comparing Hedging Performance

We follow Dumas et al. (1998) and compute hedging errors as

$$\epsilon_t = \Delta C_{t,t+h}^{actual} - \Delta C_{t,t+h}^{model}$$

which is a measure of the impact of changes in model errors from t to $t+h$ on the hedging strategy.[26] By this measure, a good model delivers hedging errors that are close to zero on average. Table V and Table VI present the results for hedging horizons from one to four weeks ahead (i.e. $h = 1, 2, 3, 4$).

Consider hedging errors at the 1-week horizon (Table Va). First, the BSM model appears to perform well, with hedging errors averaging 1.6 cents. But this hides important disparities across maturities. Average hedging errors range from 36.7 cents for out-of-the-money options to -39 cents for in-the-

[26]This abstracts from the hedging errors due to discrete adjustments. See Galai (1983) for details.

money options. Moreover, the more flexible P-BSM model has higher overall hedging errors (-4.6 cents) with substantial average errors (-18.8 cents) for the lowest strike prices.

When considering the overall mean and the dispersion of hedging errors across maturities, the best performing models are variants of the P-HG model. Identification restrictions for skewness perform well. In particular, the SP-HG2 model offers both low overall hedging errors and low dispersion across moneyness. Averages remain below 10 cents across strike prices. Table Vb draws a similar picture at the 2-week horizon. The P-BSM model sees its average performance deteriorate to -8.2 cents and mean hedging errors now range from -21.8 to 7.1 cents. Again, HG-based models offer better performance. The SP-HG2 model still offers the best performance: the mean pricing error is 0.002 cents in the entire sample and ranges from -13.6 cents to 8.6 cents across moneyness. Finally, results at the 3 and 4-week horizons (Tables (a) and (b)) quickly deteriorate for the BSM and the P-BSM models. However, the SP-HG2 model still performs well. The overall averages at 3-week and 4-week horizons are -4.3 cents and -2.1 cents.

E Discussion

Overall, the results favor the more general P-HG3 model. It offers lower in-sample and out-of-sample RMSEs as well as better hedging performances at all horizons. This contrasts with the frequent observation that the P-BSM model offers sufficient flexibility. Indeed, option prices based on the HG distribution offer better performance than the P-BSM with as many parameters (P-HG1 and P-HG2) or less (SP-HG2). If we interpret the practitioner's models as expansions around the Gaussian or the Homoscedastic Gamma distributions, the results imply that expanding around the Gaussian density is restrictive and does not offer sufficient flexibility to match the skewness and kurtosis implicit in the data. Moreover, when we consider the sequence of models, we see that imposing restrictions such that skewness is correctly measured and excess kurtosis constant does preserve most of the performance improvement.

Another way to view these results is to consider the results of Bates (2005) and Alexander and Nogueira (2005). Essentially, they show that for any contingent claim that is homogenous of degree one, all partial derivatives with respect to the underlying can be computed by taking partial derivatives of option prices with respect to strike prices. This implies that, if the number of observed option prices is arbitrarily large, we can compute delta and gamma exactly from non-parametric derivatives. In practice, however, some parametric model is fitted to observed prices from which derivatives can be imputed. The hedging performances of the P-BSM and the P-HG models imply that the latter offer a better fit of the true option price curve across the strike continuum and, therefore, a better fit of the true option's delta and gamma. In other words, the relatively poor fit of skewness by Gaussian-based expansions translates in inaccurate option sensitivity measures and larger hedging errors relative to approximations based on the Gamma density.

For our purposes, the performance of the SP-HG2 model implies that the parametric measure of risk-neutral skewness is relevant. This provides a measure of skewness that is easy to compute and requires less data than a non-parametric measure. Moreover, together with the regression results from Section IV, the importance of skewness for hedging and out-of-sample pricing confirms the key link between the risk premium and volatility shift across moneyness and skewness. Indeed, imposing the additional restriction that excess kurtosis is constant yields the next to best out-of-sample and hedging performances. Interestingly, the estimate of κ is negative (-0.042). Then relaxing the link between kurtosis and skewness allows for more asymmetry to be applied to the data than the benchmark HG model does. This adjustment is significant: to keep kurtosis constant but with κ equal to zero, skewness would have to be reduced (closer to zero) by 0.21. Taken together, the results lead us to adopt the SP-HG2 as our preferred model to measure the option-implied skewness.

VIII Term Structure Of Moments

Section V presented the trade-off between volatility and skewness when fitting option data. One important observation is that a different value of skewness was required to restore the symmetry of the IV curve for different maturities. This suggests that the risk-neutral distribution converges at slower rate than implied by the i.i.d. assumption. While the time-dependance of returns is well documented in the literature, the framework presented here allows for a transparent presentation of deviations from i.i.d. returns. We use the fact that skewness should decay toward zero with the square root of horizon, $\sqrt(H)$. If this is verified in the data, estimates of skewness multiplied by the square root of the horizon should not vary with the maturity of an options. Otherwise, the term structure of implied skewness reflects a degree of dependence implicit in option prices. Similarly, the excess kurtosis of returns should decay with H and annualized estimates of volatility should be flat across horizons. A key question is on what moment does the time dependence of returns have the greater impact.

An important advantage of our parametric approach is that we can obtain estimates of risk-neutral moments at much longer horizons than is usually the case with non-parametric methods. We estimate the term structure of volatility, skewness and kurtosis using the SP-HG2 model discussed above. We minimize pricing errors separately for each maturity (1, 2 and 3 months, and then from 4 to 6 and from 7 to 9 months. See Section III). Figure 3.9 presents the results.

Figure 3.10a presents the average (annualized) implied volatility for each maturity. The time-series average rises from close to 21.4% for the next settlement month to 21.8% at a maturity of 3 months. Thereafter, implied volatility remains more or less flat. Figure (b) presents results for (negative) the implied skewness. In contrast with implied volatility, the implied asymmetry rises sharply for all maturities we consider. Figure 3.10c shows the term structure of (negative) the implied excess kurtosis. Perhaps surprisingly, excess kurtosis relative to the HG distribution decreases with maturity. Overall, the term structure evidence indicates that the distribution of expected returns violates the i.i.d. assumptions. However, the impact of dependence appears to have a much greater impact on implied skewness than on other moments. In contrast, measures of implied volatilities flatten out beyond a maturity of 3 months while measures of implied excess kurtosis decrease with maturity. To our knowledge, this differential impact of time-dependence on skewness and kurtosis has never been documented.

IX Conclusion

We provide a simple extension of the BSM option pricing model. The Homoscedastic Gamma model allows for arbitrary skewness in the distribution of returns and delivers closed-form option pricing formula at any maturity. We provide a natural change of measure under which returns are HG under the historical and the risk-neutral probability measures. An important implication is that the relationship between the equity premium and the volatility spread is conditional on skewness. It is the ratio of the volatility spread to skewness that predicts excess returns. Empirically, we find coefficients that correspond to implications from the model. Also, the information content of the volatility spread improves when we adjust for skewness. This new stylized fact should help to discriminate among competing theories of the volatility spread.

This link between the equity premium, skewness and the volatility spread implies that skewness is key for pricing and hedging options. We first introduce the implied volatility and skewness surface, which we study empirically. This is a new tool that provide a transparent interpretation of variations in the shape and level of the IV curve in terms of skewness. Next, we develop the practitioner's version of the HG model. This approaches is robust to deviation of kurtosis from the HG model. Empirically, models based on the HG distribution perform better than their Gaussian counterparts. Hedging performances are also substantially improved. The results suggest that allowing for flexible time-variation in skewness is key for improving option pricing. Finally, we document the term structure of volatility, skewness, and kurtosis out to an horizon of 9 months. We find that dependence in returns have a larger impact on skewness than kurtosis, highlighting, again, the importance of skewness.

X Appendix

A Proposition 1

Our candidate SDF is, for given ν,

$$M_t = \exp\left(-\nu(\Delta)\varepsilon_t + \Psi(\nu(\Delta))\right),$$

where Ψ is the log-cumulant function of ε,

$$\Psi(u) = 2u\frac{\sqrt{h(\Delta)}}{\alpha(\Delta)} - \frac{4}{\alpha(\Delta)^2}\ln\left[1+\frac{1}{2}u\alpha(\Delta)\sqrt{h(\Delta)}\right].$$

Following CEFJ, this SDF defines an Equivalent Martingale Measure [EMM] if and only if

$$\Psi(\nu(\Delta)-1) - \Psi(\nu(\Delta)) - \Psi(-1) + (\mu - r_0)\Delta = 0,$$

which has the following unique solution for $\nu(\Delta)$,

$$\nu(\Delta) = -\frac{2}{\alpha(\Delta)\sqrt{h(\Delta)}} + \frac{g(\Delta)}{g(\Delta)-1},$$

where

$$g(\Delta) = \exp\left(-\frac{(\mu - r_0)\Delta}{4}\alpha(\Delta)^2 + \frac{\alpha(\Delta)\sqrt{h(\Delta)}}{2}\right).$$

Proposition 2 of CEFJ establishes sufficient conditions on Ψ for the solution to be unique.

B Limit of Risk-Neutral Volatility

Define

$$\Pi_0 \equiv (\mu - r)$$
$$\beta(\Delta) \equiv \alpha(\Delta)\frac{\sqrt{\Delta}}{2}$$
$$\sigma^*(\Delta) \equiv \sqrt{h^*(\Delta)}/\sqrt{\Delta},$$

and note that the drift correction term can be written as

$$2\frac{\sqrt{h^*(\Delta)} - \sqrt{h(\Delta)}}{\alpha(\Delta)} = \frac{\sigma^*(\Delta) - \sigma}{\beta(\Delta)}\Delta. \tag{3.17}$$

We first study the limit of the numerator as skewness tends to zero. Using the definitions above we have (see Proposition 2)

$$\sigma^*(\Delta) = \frac{g(\beta(\Delta)) - 1}{\beta(\Delta)g(\beta(\Delta))} \tag{3.18}$$

where, with a slight abuse of notation,

$$g(\beta(\Delta)) \equiv \exp(-\Pi_0\beta(\Delta)^2 + \beta(\Delta)\sigma), \tag{3.19}$$

which leads to an indeterminacy when skewness tends to zero. We use the first order expansion of the exponential function, $\exp(x) = 1 + x + x\theta(x)$ where $\theta(x)$ tends to zero when x tends to zero. Substituting in Equation 3.18 leads to, after some simplification,

$$\sigma^*(\Delta) = \frac{-\Pi_0\beta(\Delta) + \sigma + \theta(\beta(\Delta))}{1 - \Pi_0\beta(\Delta)^2 + \beta(\Delta)\sigma + \beta(\Delta)\theta(\beta(\Delta))},$$

and taking the limit shows that $\sigma^*(\Delta) \to \sigma$ when $\beta(\Delta) \to 0$.

Note then that the limit of 3.17 leads to an indeterminacy. We will again apply a Taylor expansion but, first, we compute the first order derivative of (3.18) with respect to

$\beta(\Delta)$ using Equation (3.19) to compute the derivative of $g(\beta(\Delta))$ which leads to

$$\frac{d\sigma^*(\beta(\Delta))}{d\beta(\Delta)} = \frac{1 - g(\beta(\Delta)) + \beta(\Delta)(\sigma - 2\Pi_0\beta(\Delta))}{\beta(\Delta)^2 g(\beta(\Delta))},$$

where again we face an indeterminacy. We use a second-order expansion of $g(\beta(\Delta))$

$$g(\beta(\Delta)) = g(0) + \beta(\Delta) g'(0) + \frac{1}{2} g''(0)\beta(\Delta)^2 + \beta(\Delta)^2 \theta(\beta(\Delta)),$$

where $\theta(\beta(\Delta))$ tends to zero when $\beta(\Delta)$ tends to zero. Substituting these results in a first-order expansion for $\sigma^*(\beta(\Delta))$,

$$\sigma^*(\beta(\Delta)) = \sigma^*(0) + \frac{d\sigma^*(\beta(\Delta))}{d\beta(\Delta)}(0)\,\beta(\Delta) + \beta(\Delta)\theta(\beta(\Delta)),$$

leads to

$$\frac{\sigma^*(\Delta) - \sigma}{\beta(\Delta)} = -\left(\Pi_0 + \frac{\sigma^2}{2}\right) + \theta(\beta(\Delta)),$$

which, in the limit, delivers the desired result. Note that we then have

$$\mu\Delta + 2\frac{\sqrt{h^*(\Delta)} - \sqrt{h(\Delta)}}{\alpha(\Delta} = \left(r - \frac{\sigma^2}{2}\right)\Delta + \Delta\theta(\beta(\Delta)).$$

and, finally, that if we substitute the second-order expansion for $g(\Delta)$ in the solution for ν, we get

$$\nu(\Delta) \to \frac{\mu - r + \frac{\sigma^2}{2}}{\sigma^2} = \frac{\mu - r}{\sigma^2} + \frac{1}{2},$$

C Taylor Expansion of the Price of Risk

We want to show that,

$$\nu(\beta) \approx \frac{\mu - r}{\sigma^2} + \frac{1}{2} + \frac{(\mu - r)^2 + \frac{\sigma^4}{12}}{\sigma^3}\beta$$

where

$$\begin{aligned} \nu(\beta) &= -\frac{1}{\beta\sigma} + \frac{g(\beta)}{g(\beta) - 1} \\ g(\beta) &= \exp(-(\mu - r)\beta^2 + \beta\sigma). \end{aligned}$$

Recall that $\nu(0) = (\mu - r)/\sigma^2 + \frac{1}{2}$ and note that

$$\begin{aligned} \nu'(\beta) &= \frac{1}{\beta^2\sigma} - \frac{g'(\beta)}{(g(\beta) - 1)^2} \\ g'(\beta) &= (-2(\mu - r)\beta + \sigma)\, g(\beta), \end{aligned}$$

We evaluate the limit of this derivative as $\beta \to 0$ using, as above, the second-order expansion of $g(\beta)$. After tedious but straightforward computations, the result is

$$\begin{aligned} \nu'(0) &= \frac{(\mu - r)^2 + \frac{\sigma^4}{4} - 2(\mu - r)\sigma^2 + \frac{\sigma^4}{3} - (\mu - r)\sigma^2 + 2(\mu - r)\sigma^2 + (\mu - r)\sigma^2 - \frac{\sigma^4}{2}}{\sigma^3} \\ &= \frac{(\mu - r)^2 + \frac{\sigma^4}{12}}{\sigma^3}. \end{aligned}$$

D Proposition 2

From CEFJ, the logarithm risk-neutral of the risk-neutral Moment Generating Function is

$$\begin{aligned} \Psi^{Q*}(u) &= -u\Psi'(\nu(\Delta)) + \Psi(\nu(\Delta) + u) - \Psi(\nu(\Delta)) \\ &= 2u\frac{\sqrt{h^*(\Delta)}}{\alpha(\Delta)} - \frac{4}{\alpha(\Delta)^2}\ln\left[1 + \frac{1}{2}u\alpha(\Delta)\sqrt{h^*(\Delta)}\right], \end{aligned}$$

implying that

$$\varepsilon^*_{t+\Delta} = \frac{\sqrt{h(\Delta)}}{\sqrt{h^*(\Delta)}}\varepsilon_{t+\Delta} + \nu(\Delta)\sqrt{h(\Delta)}.$$

The HG model can then be written as

$$\ln(S_{t+\Delta}/S_t) = r_0\Delta - \gamma^*(\Delta) + \sqrt{h^*(\Delta)}\varepsilon^*_{t+\Delta},$$

where

$$\gamma^*(\Delta) = \Psi^{Q*}(-1) = -2\frac{\sqrt{h^*(\Delta)}}{\alpha(\Delta)} - \frac{4}{\alpha(\Delta)^2}\ln\left[1 - \frac{1}{2}\alpha(\Delta)\sqrt{h^*(\Delta)}\right],$$

and with

$$\sqrt{h^*(\Delta)} = \frac{2(g(\Delta)-1)}{\alpha(\Delta)\,g(\Delta)}.$$

Substituting back in the equation for returns under the risk-neutral measure, and simplifying, yields the results.

E Greeks

For notational simplicity we introduce $a \equiv H/\beta(\Delta)^2$. We begin with the sensitivity to changes in the underlying stock price. The HG option price is homogenous of degree one in stock price and strike. Then the standard result holds and the option delta is simply

$$\frac{\partial C_t}{\partial S_t} = C_{1,t}, \tag{3.20}$$

which depends on skewness. Next, the sensitivity of the option's delta with respect to the stock price is

$$\frac{\partial^2 C_t}{\partial S_t^2} = \frac{e^{-(d_2+r_f H)}d_2^{a-1}}{|\beta|\sigma^*\Gamma(a)}\frac{K}{S_t^2}, \tag{3.21}$$

which also depends on skewness and moneyness. The sensitivity of option prices to changes in the underlying risk-neutral volatility is

$$\frac{\partial C_t}{\partial \sigma_t^*} = \frac{|\beta|\sigma^* e^{(-r_f H)}K}{\sigma^*(1-\beta\sigma^*)}\frac{e^{-d_2}d_2^a}{\Gamma(a)}, \tag{3.22}$$

and, finally, the sensitivity of option prices to changes in the skewness of returns is given by

$$\begin{aligned}
\frac{\partial C_t}{\partial \beta} &= -\frac{2a}{\beta}\left[(\ln(d_2) - \Psi(a))\,C_t - Ke^{(-r_f H)}P(a,d_2)\ln(1-\beta\sigma)\right] \quad (3.23)\\
&+ \frac{2a}{\beta}\Gamma(a)d_2^a S_t(1-\beta\sigma)^a{}_2\bar{F}_2(a,a;a+1,a+1;-d_1)\\
&- \frac{2a}{\beta}\Gamma(a)d_2^a Ke_2^{(-r_f H)}\bar{F}_2(a,a;a+1,a+1;-d_2)\\
&- Ke^{(-r_f H)}\frac{\sigma^*}{1-\beta\sigma^*}\frac{e^{-d_2}d_2^a}{\Gamma(a)},
\end{aligned}$$

where

$$\Psi(a,z) = P(a,z)\ln(z) - \Gamma(a)\,z^a{}_2\bar{F}_2(a,a;a+1,a+1;-z),$$

and where ${}_2\bar{F}_2(\cdot)$ is the regularized hypergeometric function.

F Proposition 3

A no-arbitrage price of a European call option with strike price K and maturity T can be obtained from the computation of the discounted expectation of the terminal payoff under the risk-neutral measure. That is,

$$\begin{aligned}
C_t(K,M) &= E^Q[\max(S_{t+T}-K,0)]\\
C_t &= \exp(-rT)S_t E^Q\left[\exp(R_{t,M})\,1_{[R_{t,M}>\ln(K/S_t)]}\right] - \exp(-r_0T)\,KP^Q\left[R_{t,M} > \ln(K/S_t)\right].
\end{aligned}$$

We can compute $P^Q[R_{t,M} > \ln(K/S_t)]$ from the distribution function of a gamma variable. Note first that

$$P^Q[R_{t,M} > \ln(K/S_t)] = P^Q\left[\frac{\beta(\Delta)}{\sqrt{\Delta M}} y^*_{t,M} > \frac{\ln(K/S_t) - \mu^*(\Delta) M\Delta}{\sqrt{\Delta M}\sigma^*(\Delta)} + \frac{\sqrt{\Delta M}}{\beta(\Delta)}\right],$$

where we define

$$\frac{2\sqrt{M}}{\alpha(\Delta)}\left(\varepsilon^*_{t,M} + \frac{2\sqrt{M}}{\alpha(\Delta)}\right) = y^*_{t,M} \sim^Q \Gamma\left(\frac{4M}{\alpha(\Delta)^2}, 1\right),$$

based on the characterization of the standardized Gamma distribution given in Equation 3.2. If $\alpha(\Delta) > 0$,

$$P^Q[R_{t,M} > \ln(K/S_t)] == \frac{\Gamma\left(\frac{T}{\beta^2(\Delta)}; \frac{T}{\beta^2(\Delta)} + \frac{\ln(K/S_t) - \mu^*(\Delta)T}{\beta(\Delta)\sigma^*(\Delta)}\right)}{\Gamma\left(\frac{T}{\beta^2(\Delta)}\right)},$$

where $\Gamma(a, x)$ is the upper incomplete gamma function[27] and if $\alpha(\Delta) < 0$,

$$\begin{aligned} P^Q[R_{t,M} > \ln(K/S_t)] &= = \frac{\gamma\left(\frac{T}{\beta(\Delta)^2}; \frac{T}{\beta(\Delta)^2} + \frac{\ln(K/S_t) - \mu^* T}{\beta(\Delta)\sigma^*(\Delta)}\right)}{\Gamma\left(\frac{T}{\beta(\Delta)^2}\right)} \\ &= 1 - \frac{\Gamma\left(\frac{T}{\beta(\Delta)^2}; \frac{T}{\beta(\Delta)^2} + \frac{\ln(K/S_t) - \mu^* T}{\beta(\Delta)\sigma^*(\Delta)}\right)}{\Gamma\left(\frac{T}{\beta(\Delta)^2}\right)}. \end{aligned}$$

Similarly,

$$\begin{aligned} & E^Q\left[\exp(R_{t,M})\, 1_{[R_{t,M} > \ln(K/S_t)]}\right] \\ = \; & \exp\left(\mu^*(\Delta) M\Delta - \frac{\sigma^*(\Delta) M\Delta}{\beta(\Delta)}\right) E^Q\left[\exp\left(\sigma^*(\Delta)\beta(\Delta) y^*_{t,M}\right) 1_{\left[\frac{\beta(\Delta)}{\sqrt{\Delta M}} y^*_{t,M} > \kappa\right]}\right] \end{aligned}$$

where we use

$$\kappa = \frac{\ln(K/S_t) - \mu^*(\Delta) M\Delta}{\sqrt{\Delta M}\sigma^*(\Delta)} + \frac{\sqrt{\Delta M}}{\beta(\Delta)}.$$

Then, if $\alpha(\Delta) > 0$, and using that $y^*_{t,M}$ has a standard gamma distribution with parameter $\frac{M\Delta}{\beta(\Delta)^2}$, we have

$$\begin{aligned} & E^Q\left[\exp\left(\sigma^*(\Delta)\beta(\Delta) y^*_{t,M}\right) 1_{\left[y^*_{t,M} > \frac{\sqrt{\Delta M}\kappa}{\beta(\Delta)}\right]}\right] \\ = \; & \int_{(1-\sigma^*(\Delta)\beta(\Delta))\frac{\sqrt{\Delta M}\kappa}{\beta(\Delta)}}^{\infty} \exp\left(-z^*_{t,M}\right) \frac{\left(z^*_{t,M}\right)^{\frac{M\Delta}{\beta(\Delta)^2}-1}}{\left(1 - \sigma^*(\Delta)\beta(\Delta)\right)^{\frac{M\Delta}{\beta(\Delta)^2}} \Gamma\left(\frac{M\Delta}{\beta(\Delta)^2}\right)} dz^*_{t,M} \\ = \; & \frac{\Gamma\left(\frac{M\Delta}{\beta(\Delta)^2}; \left(\frac{M\Delta}{\beta(\Delta)^2} + \frac{\ln(K/S_t) - \mu^*(\Delta)\Delta M}{\beta(\Delta)\sigma^*(\Delta)}\right)\left(1 - \sigma^*(\Delta)\beta(\Delta)\right)\right)}{\Gamma\left(\frac{M\Delta}{\beta(\Delta)^2}\right)\left(1 - \sigma^*(\Delta)\beta(\Delta)\right)^{\frac{M\Delta}{\beta(\Delta)^2}}} \end{aligned}$$

and, using the change of variables $(1 - \sigma^*(\Delta)\beta(\Delta)) y^*_{t,M} = z^*_{t,M}$, it follows that

$$\begin{aligned} & E^Q\left[\exp(R_{t,M})\, 1_{[R_{t,M} > \ln(K/S_t)]}\right] \\ = \; & \exp\left(\left(\mu^*(\Delta) - \frac{\sigma^*(\Delta)}{\beta(\Delta)}\right) T\right) \frac{\Gamma\left(\frac{T}{\beta(\Delta)^2}; \left(\frac{T}{\beta(\Delta)^2} + \frac{\ln(K/S_t) - \mu^*(\Delta)T}{\beta(\Delta)\sigma^*(\Delta)}\right)\left(1 - \sigma^*(\Delta)\beta(\Delta)\right)\right)}{\Gamma\left(\frac{T}{\beta(\Delta)^2}\right)\left(1 - \sigma^*(\Delta)\beta(\Delta)\right)^{\frac{T}{\beta(\Delta)^2}}}. \end{aligned}$$

[27]The upper incomplete gamma function is defined as $\Gamma(a, x) = \int_x^\infty t^{a-1} e^{-t} dt$ while the lower incomplete gamma function is defined as $\gamma(a, x) = \int_0^x t^{a-1} e^{-t} dt$. Note that $\Gamma(a) = \Gamma(a, 0)$ while $\gamma(a) = \gamma(a, \infty)$.

If, however, $\alpha(\Delta) < 0$ then

$$E^Q\left[\exp\left(\sigma^*(\Delta)\beta(\Delta)y^*_{t,M}\right)1_{\left[\frac{\beta(\Delta)}{\sqrt{\Delta M}}y^*_{t,M}>\kappa\right]}\right]$$
$$=\frac{\gamma\left(\frac{M\Delta}{\beta(\Delta)^2};\left(\frac{M\Delta}{\beta(\Delta)^2}+\frac{\ln(K/S_t)-\mu^*(\Delta)\Delta M}{\beta(\Delta)\sigma^*(\Delta)}\right)(1-\sigma^*(\Delta)\beta(\Delta))\right)}{\Gamma\left(\frac{M\Delta}{\beta(\Delta)^2}\right)(1-\sigma^*(\Delta)\beta(\Delta))^{\frac{M\Delta}{\beta(\Delta)^2}}},$$

and then

$$E^Q\left[\exp(R_{t,M})1_{[R_{t,M}>\ln(K/S_t)]}\right]$$
$$=\exp\left(\left(\mu^*(\Delta)-\frac{\sigma^*(\Delta)}{\beta(\Delta)}\right)T\right)\frac{\gamma\left(\frac{T}{\beta(\Delta)^2};\left(\frac{T}{\beta(\Delta)^2}+\frac{\ln(K/S_t)-\mu^*(\Delta)T}{\beta(\Delta)\sigma^*(\Delta)}\right)(1-\sigma^*(\Delta)\beta(\Delta))\right)}{\Gamma\left(\frac{T}{\beta(\Delta)^2}\right)(1-\sigma^*(\Delta)\beta(\Delta))^{\frac{T}{\beta(\Delta)^2}}}.$$

G Proposition 4

Suppose that the underlying stock price evolution under the risk-neutral measure is given by

$$R_T = (r-\delta)T + \sigma\sqrt{T}y$$

where δ is a risk-adjustment factor, y is a random number with mean zero, variance 1, skewness, $\frac{2\beta}{\sqrt{T}}$ and kurtosis, λ_2. Suppose also that the probability density of y is described by the following Edgeworth series expansion around the standardized gamma distribution:

$$f(y) = g(y) + \frac{\lambda_2 - \frac{6\beta^2}{T}}{4!}\frac{d^4g(y)}{dy^4},$$

where $g(y)$ is the standardized gamma density function given by

$$g(y) = \frac{\sqrt{T}z^{a-1}e^{-z}}{|\beta|\,\Gamma(a)} \quad if \;\; \beta y > -\sqrt{T},$$

and where $z = \frac{\sqrt{T}}{\beta}y + a$. Imposing that gross stock returns are a martingale under the risk-neutral measure,

$$\begin{aligned} E_0^Q[\exp(R_T)] &= E_0^Q[\exp\left((r-\delta)T+\sigma\sqrt{T}y\right)] \\ &= \exp((r-\delta)T)\int\exp\left(\sigma\sqrt{T}y\right)\left[g(y)+\frac{\lambda_2-\frac{6\beta^2}{T}}{4!}\frac{d^4g(y)}{dy^4}\right]dy, \end{aligned}$$

leads to the required risk-adjustment,

$$\delta = \frac{1}{T}\ln\left[\exp\left(-\frac{\sigma T}{\beta}-a\ln(1-\beta\sigma)\right)+\frac{\lambda_2-\frac{6\beta^2}{T}}{4!}\int\exp\left(\sigma\sqrt{T}y\right)\frac{d^4g(y)}{dy^4}dy\right].$$

The price of a European call option is

$$c_0^* = e^{-rT}\int_{-d_2^*}^{\infty}\left(S_0\exp\left((r-\delta)T+\sigma\sqrt{T}y\right)-K\right)f(y)\,dy$$

where

$$d_2^* = \frac{\ln(S_0/K)+(r-\delta)T}{\sigma\sqrt{T}}.$$

We have

$$c_0^* = e^{-rT}\left[S_0\exp((r-\delta)T)\int_{-d_2^*}^{\infty}\exp\left(\sigma\sqrt{T}y\right)f(y)\,dy - K\int_{-d_2^*}^{\infty}f(y)\,dy\right].$$

For the first integral, we have

$$\int_{-d_2^*}^{\infty} \exp\left(\sigma\sqrt{T}y\right) f(y)\, dy = \int_{-d_2^*}^{\infty} \exp\left(\sigma\sqrt{T}y\right) g(y)\, dy + \kappa \int_{-d_2^*}^{\infty} \exp\left(\sigma\sqrt{T}y\right) \frac{d^4 g(y)}{dy^4} dy$$

and for $\beta \leq 0$, say, and $d_1^* = \bar{d}_2 (1 - \sigma\beta)$ we have

$$\int_{-d_2^*}^{\infty} \exp\left(\sigma\sqrt{T}y\right) g(y)\, dy = \frac{e^{-\sigma\beta a}}{(1-\sigma\beta)^a} P(a, d_1^*)$$

while

$$\begin{aligned} & \int_{-d_2^*}^{\infty} \exp\left(\sigma\sqrt{T}y\right) \frac{d^4 g(y)}{dy^4} dy \\ = \quad & a^2 e^{-\sigma\beta a} \left[\begin{array}{c} \frac{P(a-4,d_1^*)}{(1-\sigma\beta)^{a-4}} - 4\frac{P(a-3,d_1^*)}{(1-\sigma\beta)^{a-3}} \\ +6\frac{P(a-2,d_1^*)}{(1-\sigma\beta)^{a-2}} - 4\frac{P(a-1,d_1^*)}{(1-\sigma\beta)^{a-1}} + \frac{P(a,d_1^*)}{(1-\sigma\beta)^{a}} \end{array} \right]. \end{aligned}$$

Next, for the second integral above,

$$\int_{-d_2^*}^{\infty} f(y)\, dy = \int_{-d_2^*}^{\infty} g(y)\, dy + \kappa \int_{-d_2^*}^{\infty} \frac{d^4 g(y)}{dy^4} dy$$

with

$$\begin{aligned} \int_{-d_2^*}^{\infty} g(y)\, dy &= \int_0^{a - d_2^* \frac{\sqrt{T}}{\beta}} \frac{z^{a-1} e^{-z}}{\Gamma(a)} dz = P(a, \bar{d}_2) \\ \int_{-d_2^*}^{\infty} \frac{d^4 g(y)}{dy^4} dy &= a^2 \left[\begin{array}{c} P(a-4, \bar{d}_2) - 4P(a-3, \bar{d}_2) + \\ 6P(a-2, \bar{d}_2) - 4P(a-1, \bar{d}_2) + P(a, \bar{d}_2) \end{array} \right]. \end{aligned}$$

H Identifying Restriction on the P-HG

The equality of prices from the true model and the P-HG for at-the-money options implies that

$$\begin{aligned} P(a, d_1^*) - P(a, d_2^*) = {} & P(a, d_1) - \left(1 + T^2\sigma^4\kappa\right) P(a, d_2) \\ & + \kappa \frac{T^2\sigma}{\beta^3} \left[-h''(d_2) + \sigma\beta h'(d_2) - \sigma^2\beta^2 h(d_2) \right], \end{aligned}$$

while the equality of the first derivative of prices implies

$$\begin{aligned} P(a, d_2^*) + \frac{\sigma_{I0}\gamma_1}{\bar{\sigma}\sqrt{T}} \frac{d_2^*\beta h(d_2^*)}{(1-\beta\sigma_{0I})} = {} & \left(1 + T^2\sigma^4\kappa\right) P(a, d_2) \\ & + \kappa a^2 \left[h'''(d_2) + \sigma^3\beta^3 h(d_2) \right], \end{aligned}$$

and, finally, the equality of the second derivatives implies

$$\begin{aligned} & \frac{h(d_2^*)}{\sigma_{0I}} \left[1 + \frac{(2a - d_1^* - d_2^*)\beta\sigma_{I0}\gamma_1}{(1-\beta\sigma_{0I})\bar{\sigma}\sqrt{T}} + \frac{d_2^*\beta^3\sigma_{I0}^3\gamma_1^2}{(1-\beta\sigma_{0I})^2\bar{\sigma}^2 T} + \frac{2d_2^*\beta^2\sigma_{I0}^2\gamma_2}{(1-\beta\sigma_{0I})\bar{\sigma}^2 T} \right] = \\ & \left(1 + T^2\sigma^4\kappa\right) \frac{h(d_2)}{\sigma} + \frac{\kappa a^2}{\sigma} \left[h^{(4)}(d_2) + \sigma^3\beta.^3 h'(d_2) \right]. \end{aligned}$$

Then, linearizing the left sides of the equations around $\sigma_0 = \sigma$, $\gamma_1 = 0$ and $\gamma_2 = 0$, respectively, and the right side around $\kappa = 0$ leads to

$$\begin{aligned}
\frac{\sigma_{I0} - \sigma}{\sigma} &= a^2 (1 - \sigma\beta) \left(\sigma^3 \beta^3 \frac{P(a, d_2)}{h(d_2) d_2} + \frac{(a-1)(a-2)}{d_2^3} - \frac{(a-1)(2 + \sigma\beta)}{d_2^2} + \frac{1 + \sigma\beta + \sigma^2\beta^2}{d_2} \right) \kappa \\
\gamma_1 &= -\frac{\bar{\sigma}\sqrt{T}(a - d_1)}{\beta\sigma_2} \frac{\sigma_{I0} - \sigma}{\sigma} + \frac{\bar{\sigma}\sqrt{T} a^2 (1 - \sigma\beta)}{d_2} \left[\frac{\beta^3 \sigma^3 P(a, d_2)}{h(d_2)} + 2\beta^2\sigma^2 + \frac{h^{(3)}(d_2)}{\beta\sigma h(d_2)} \right] \kappa \\
\gamma_2 &= -\frac{\bar{\sigma}^2 T}{2\beta^2\sigma^2 d_2} \left(\frac{h'(d_2)(a - d_1)}{h(d_2)} - 1 + \sigma\beta \right) \frac{\sigma_{I0} - \sigma}{\sigma} \\
&\quad + - \frac{\bar{\sigma}\sqrt{T}(2a - d_1 - d_2)}{2 d_2 \beta\sigma} \gamma_1 + \left(\sigma + \frac{h'(d_2)}{\beta h(d_2)} \right) \frac{\sigma(1 - \sigma\beta)\bar{\sigma}^2 T^3}{2 d_2 \beta^2} \kappa
\end{aligned}$$

where

$$\begin{aligned}
d_2 &= \frac{-a \ln(1 - \sigma\beta)}{\sigma\beta} \\
d_1 &= d_2 (1 - \sigma\beta) \\
a &= \frac{T}{\beta^2} \\
\kappa &= \frac{\lambda_2 - \frac{6\beta^2}{T}}{4!}.
\end{aligned}$$

Bibliography

Alexander, C. and L. M. Nogueira (2005). Model-free hedge ratios and scale-invariant models. *Journal of Banking and Finance 31*, 1839–1861.

Andersen, T., T. Bollerslev, and F. Diebold (2005). Parametric and nonparametric volatility measurement. forthcoming in *The Handbook of Financial Econometrics*.

Bakshi, G., C. Cao, and Z. Chen (1997). Empirical performance of alternative option pricing models. *The Journal of Finance 51*, 549–584.

Bakshi, G. and N. Kapadia (2003). Delta-hedged gains and the negative market volatility risk premium. *Review of Financial Studies 16*, 527–566.

Bakshi, G., N. Kapadia, and D. Madan (2003). Stock return characteristics, skew laws, and the differential pricing of individual equity options. *Review of Financial Studies 16*, 101–143.

Bakshi, G. and D. Madan (2000). Spanning and derivative-security valuation. *Journal of Financial Economics 58*, 205–238.

Bakshi, G. and D. Madan (2006). A theory of volatility spreads. *Management Science 52*, 1945–1956.

Bates, D. (2000). Post-'87 crash fears in the SP500 futures option market. *Journal of Econometrics 94*, 181–238.

Bates, D. (2005). Hedging the smirk. *Financial Research Letters 2*, 195–200.

Bates, D. S. (1995). Testing option pricing models. Working Paper 5129, National Bureau of Economic Research.

Bollerslev, T., G. Tauchen, and H. Zhou (2008). Expected stock returns and variance risk premia. *ERID Working Paper Series 5.*

Carr, P. and L. Wu (2009). Variance risk premiums. *Review of Financial Studies 22.*

Christoffersen, P., R. Elkamhi, B. Feunou, and K. Jacobs (2009). Option valuation with conditional heteroskedasticity and non-normality. *Review of Financial Studies*. forthcoming.

Christoffersen, P., S. Heston, and K. Jacobs (2006). Option valuation with conditional skewness. *Journal of Econometrics 131*, 253–284.

Christoffersen, P., K. Jacobs, and G. Vainberg (2008). Forward-looking betas.

Corrado, C. and T. Su (1996). Skewness and kurtosis in SP500 index returns implied by options prices. *The Journal of Financial Research 19.*

Dennis, P. and S. Mayhew (2000). Risk-neutral skewness: Evidence from stock options. *The Journal of Financial and Quantitative Analysis 37*, 471–493.

Dumas, B., J. Fleming, and R. Whaley (1998). Implied volatility functions: Empirical tests. *The Journal of Finance 53*, 2059–2106.

Galai, D. (1983). The components of the returns from hedging options against stocks. *The Journal of Business 56*, 45–54.

Jarrow, R. and A. Rudd (1982). Approximate option valuation for arbitrary stochastic processes. *Journal of Financial Economics 10.*

Jondeau, E. and M. Rockinger (2001). Gram-charlier densities. *Journal of Economic Dynamics & Control 25.*

Kim, T.-H. and H. White (2003). On more robust estimation of skewness and kurtosis: simulation and application to the SP500 index.

León, A., J. Mencía, and E. Sentana (2006). Parametric propertis of semi-nonparametric distributions, with applications to option valuation. CEMFI.

Polimenis, V. (2006). Skewness corrections for asset pricing. Working Paper.

Potters, M., R. Cont, and J.-P. Bouchaud (1998). Financial markets as adaptative systems. *Europhysics Letters 41.*

Rompolis, L. and E. Tzavalis (2008). The effects of the risk-neutral skewness on implied volatility regressions. Working Paper.

Rubinstein, M. and J. Jackwerth (1998). Recovering probability distribution from option prices. *The Journal of Finance 51*, 1611–1631.

Zhang, J. and Y. Xiang (2005). Implied volatility smirk. Working Paper.

Table I: Summary statistics for strike price and maturity categories.

(a) Summary statistics by moneyness

	Moneyness					
	<0.95	<0.975	<1	<1.025	>1.025	All
Number of Contracts	3343	2418	3859	3077	3809	16506
Average Call Price	28.24	31.80	37.22	47.05	78.85	46.05
Average IV	19.43	19.23	19.36	20.13	22.66	20.26

(b) Summary statistics by maturities

	Contract Month						
	1	2	3	4-6	7-9	10-12	All
Number of Contracts	4303	4016	2377	2822	1726	1167	16506
Average Call Price	36.60	39.53	42.91	51.53	61.95	72.74	46.05
Average IV	20.47	20.24	20.37	20.19	20.15	20.24	20.26

(c) Summary statistics by moneyness and maturities. For each moneyness and strike price category, the first line gives the number of contracts and the second line gives the average Implied Volatility.

	Moneyness				
Months	<0.95	0.95 to 0.975	0.975 to 1	1 to 1.025	>1.025
1	96	398	1104	1172	1533
	21.39	18.65	18.63	19.55	22.92
2	354	668	1113	848	1033
	19.80	18.66	19.13	20.08	22.75
3	461	445	647	406	418
	19.75	19.24	19.78	20.94	22.61
4-6	973	481	504	371	493
	19.27	19.48	20.00	20.88	22.39
7-9	805	262	280	167	212
	19.18	20.35	20.33	21.26	22.46
10-12	639	157	194	89	88
	19.44	20.72	20.99	21.48	22.30

Table II: **Predictability of Excess Returns by Implied Skewness.**

The table reports the results of n-period regressions of returns on the SP500 index in excess of a yield of maturity of n months:

$$\frac{1}{n}\sum_{j=1}^{n}\left(r_{M,t+j} - y_{f,t+j}^{(n)} + \frac{IV_t}{2}\right) = a_n + b_n^{\top} PRED_t + \varepsilon_{n,t+n}.$$

The regressor PRED is a combination of IV-RV and (IV-RV)/IS, where IV and IS are annualized implied volatility and skewness from all option contracts, and RV is the annualized realized volatility. Reported in square brackets and in brackets are respective robust t-statistics for the null that the coefficient is equal to zero, and for the null that the coefficient is equal to −2. The sample period is from January 1996 to December 2004.

	1	3	6	12	24	36
Constant	-22.19	-5.43	-3.50	-7.14	-6.93	-18.96
	[-0.65]	[-0.20]	[-0.12]	[-0.24]	[-0.24]	[-0.70]
(IV-RV)/IS	-3.28	-2.24	-2.04	-2.13	-1.58	-1.64
	[-2.66]	[-2.52]	[-2.69]	[-3.85]	[-2.38]	[-2.66]
	(-1.04)	(-0.27)	(-0.05)	(-0.23)	(0.64)	(0.57)
Adj. R^2	1.85	3.11	5.59	9.72	8.06	11.28
Constant	0.10	2.86	-8.13	-10.68	-0.63	2.31
	[0.00]	[0.08]	[-0.26]	[-0.33]	[-0.02]	[0.07]
IV-RV	7.33	6.38	8.11	8.28	4.37	2.12
	[1.76]	[1.65]	[3.01]	[3.40]	[1.51]	[0.75]
Adj. R^2	-0.03	1.18	5.83	9.72	3.52	-0.11
Constant	-11.78	-3.23	-10.59	-13.93	-5.15	-7.55
	[-0.33]	[-0.10]	[-0.34]	[-0.44]	[-0.16]	[-0.25]
IV-RV	-7.46	-1.53	4.83	4.63	-1.15	-5.29
	[-0.93]	[-0.25]	[1.18]	[1.24]	[-0.27]	[-1.71]
(IV-RV)/IS	-4.79	-2.55	-1.06	-1.19	-1.81	-2.59
	[-1.98]	[-1.66]	[-0.86]	[-1.35]	[-1.98]	[-3.31]
Adj. R^2	1.27	2.21	5.55	10.05	7.05	14.06

Table III: In-sample RMSE

RMSE by moneyness and by maturity in percentage of BSM model's RMSE. BSM is the Black-Scholes Model, HG is the Homoscedastic Gamma Model, P-BSM and P-HG are practitioner's versions of these models where volatility is quadratic in moneyness. P-HG1 is a version where the linear term is zero (i.e. $\gamma_1 = 0$), P-HG2 imposes that β is the risk-neutral skewness (see text) and P-HG3 is unrestricted. SP-BSM and SP-HG are smoothed version of these models where the shape of the quadratic IV curve is constant through the sample.

(a) In-sample RMSE by moneyness

Model			**Moneyness**			
	S/X<0.95	**0.95<S/X<0.975**	**0.975<S/X<1**	**1<S/X<1.025**	**1.025<S/X**	**All**
HG	0.584	0.639	0.649	0.681	0.570	0.368
P-BSM	0.487	0.829	0.901	0.665	0.351	0.350
P-HG1	0.437	0.537	0.536	0.629	0.595	0.278
P-HG2	0.449	0.583	0.579	0.532	0.410	0.234
P-HG3	0.298	0.473	0.471	0.453	0.313	0.135
SP-BSM	0.545	0.812	0.919	0.759	0.477	0.418
SP-HG1	0.562	0.632	0.709	0.712	0.655	0.399
SP-HG2	0.488	0.629	0.709	0.642	0.505	0.312
SP-HG3	0.485	0.667	0.744	0.692	0.469	0.322

(b) In-sample RMSE by maturity

Model			**Maturity**			
	1	**2**	**3**	**4-6**	**7-9**	**All**
HG	0.916	0.729	0.585	0.385	0.569	0.368
P-BSM	0.891	0.735	0.614	0.382	0.526	0.350
P-HG1	0.845	0.697	0.544	0.355	0.420	0.278
P-HG2	0.652	0.593	0.529	0.335	0.437	0.234
P-HG3	0.547	0.450	0.405	0.290	0.305	0.135
SP-BSM	1.018	0.798	0.624	0.379	0.592	0.418
SP-HG1	0.892	0.800	0.669	0.493	0.530	0.399
SP-HG2	0.771	0.647	0.612	0.438	0.505	0.312
SP-HG3	0.916	0.678	0.542	0.335	0.525	0.322

Table IV: Out-of-sample RMSE

Weekly out-of-sample RMSE by moneyness and by maturity in percentage of BSM model's RMSE. Parameters obtained for a given week are held constant to price options observed the following week. BSM is the Black-Scholes Model, HG is the Homoscedastic Gamma Model, P-BSM and P-HG are practitioner's versions of these models where volatility is quadratic in moneyness. P-HG1 is a version where the linear term is zero (i.e. $\gamma_1 = 0$), P-HG2 imposes that β is the risk-neutral skewness (see text) and P-HG3 is unrestricted. SP-BSM and SP-HG are smoothed version of these models where the shape of the quadratic IV curve is constant through the sample.

(a) Out-of-sample RMSE by moneyness

Model			**Moneyness**			
	S/X<0.95	**0.95<S/X<0.975**	**0.975<S/X<1**	**1<S/X<1.025**	**1.025<S/X**	**All**
HG	0.795	0.895	0.877	0.840	0.715	0.657
P-BSM	0.718	0.936	0.999	0.914	0.748	0.676
P-HG1	0.736	0.888	0.869	0.833	0.737	0.621
P-HG2	0.730	0.906	0.892	0.829	0.840	0.658
P-HG3	0.656	0.855	0.876	0.832	0.733	0.568
SP-BSM	0.745	0.930	0.999	0.908	0.683	0.671
SP-HG1	0.774	0.880	0.904	0.860	0.763	0.665
SP-HG2	0.724	0.865	0.895	0.852	0.801	0.639
SP-HG3	0.725	0.885	0.927	0.872	0.696	0.625

(b) Out-of-Sample RMSE by maturity

Model			**Maturity**			
	1	**2**	**3**	**4-6**	**7-9**	**All**
HG	1.059	0.894	0.859	0.715	0.757	0.657
P-BSM	1.069	0.942	0.886	0.727	0.744	0.676
P-HG1	1.023	0.902	0.871	0.718	0.695	0.621
P-HG2	1.264	0.914	0.876	0.708	0.695	0.658
P-HG3	0.996	0.871	0.865	0.717	0.628	0.568
SP-BSM	1.068	0.923	0.864	0.708	0.764	0.671
SP-HG1	1.002	0.933	0.896	0.756	0.728	0.665
SP-HG2	1.055	0.894	0.857	0.725	0.724	0.639
SP-HG3	1.040	0.877	0.855	0.702	0.726	0.625

Table V: Hedging Errors I

Weekly hedging errors by moneyness in dollars. BSM is the Black-Scholes Model, HG is the Homoscedastic Gamma Model, P-BSM and P-HG are practitioner's versions of these models where volatility is quadratic in moneyness. P-HG1 is a version where the linear term is zero (i.e. $\gamma_1 = 0$), P-HG2 imposes that β is the risk-neutral skewness (see text) and P-HG3 is unrestricted. SP-BSM and SP-HG are smoothed version of these models where the shape of the quadratic IV curve is constant through the sample.

(a) 1-week Hedging Horizon

Model	**Moneyness**					
	S/X<0.95	**0.95<S/X<0.975**	**0.975<S/X<1**	**1<S/X<1.025**	**1.025<S/X**	**All**
BSM	0.367	0.200	-0.031	-0.207	-0.390	0.016
HG	0.141	-0.204	-0.212	-0.021	0.177	-0.035
P-BSM	-0.188	-0.124	-0.022	0.127	0.021	-0.046
P-HG1	0.072	-0.103	-0.094	0.042	0.127	0.001
P-HG2	0.085	-0.117	-0.136	0.132	0.174	0.014
P-HG3	-0.028	-0.105	-0.048	0.101	0.135	0.001
SP-BSM	-0.123	-0.122	-0.070	0.048	0.039	-0.054
SP-HG1	-0.154	-0.071	0.086	0.260	0.025	0.023
SP-HG2	0.075	-0.077	-0.096	-0.003	0.024	-0.018
SP-HG3	-0.041	-0.146	-0.112	0.004	0.070	-0.053

(b) 2-week Hedging Horizon

Model	**Moneyness**					
	S/X<0.95	**0.95<S/X<0.975**	**0.975<S/X<1**	**1<S/X<1.025**	**1.025<S/X**	**All**
BSM	0.546	0.219	-0.059	-0.429	-0.817	-0.019
HG	0.182	-0.311	-0.239	-0.076	0.013	-0.082
P-BSM	-0.219	-0.122	0.018	0.071	-0.130	-0.082
P-HG1	0.047	-0.129	-0.043	0.060	-0.039	-0.019
P-HG2	0.131	-0.122	-0.070	0.174	0.168	0.046
P-HG3	-0.030	-0.090	0.003	0.131	0.030	0.002
SP-BSM	-0.176	-0.168	-0.059	0.006	-0.161	-0.114
SP-HG1	-0.231	-0.021	0.252	0.363	-0.138	0.037
SP-HG2	0.086	-0.068	0.034	0.027	-0.136	0.002
SP-HG3	-0.082	-0.250	-0.112	-0.019	-0.049	-0.106

Table VI: Hedging Errors II

Weekly hedging errors by moneyness in dollars. BSM is the Black-Scholes Model, HG is the Homoscedastic Gamma Model, P-BSM and P-HG are practitioner's versions of these models where volatility is quadratic in moneyness. P-HG1 is a version where the linear term is zero (i.e. $\gamma_1 = 0$), P-HG2 imposes that β is the risk-neutral skewness (see text) and P-HG3 is unrestricted. SP-BSM and SP-HG are smoothed version of these models where the shape of the quadratic IV curve is constant through the sample.

(a) 3-week Hedging Horizon

Model	**Moneyness**					
	S/X<0.95	**0.95<S/X<0.975**	**0.975<S/X<1**	**1<S/X<1.025**	**1.025<S/X**	**All**
BSM	0.707	0.383	-0.134	-0.652	-1.212	0.015
HG	0.211	-0.341	-0.345	-0.150	-0.071	-0.116
P-BSM	-0.330	-0.189	-0.111	-0.081	-0.184	-0.197
P-HG1	0.067	-0.120	-0.068	0.024	-0.119	-0.031
P-HG2	0.095	-0.110	-0.062	0.185	0.155	0.037
P-HG3	-0.073	-0.108	0.014	0.066	0.013	-0.029
SP-BSM	-0.315	-0.248	-0.176	-0.067	-0.233	-0.223
SP-HG1	-0.307	0.049	0.325	0.379	-0.250	0.018
SP-HG2	-0.002	-0.024	-0.020	-0.035	-0.233	-0.043
SP-BS3	-0.227	-0.331	-0.216	-0.069	-0.134	-0.211

(b) 4-week Hedging Horizon

Model	**Moneyness**					
	S/X<0.95	**0.95<S/X<0.975**	**0.975<S/X<1**	**1<S/X<1.025**	**1.025<S/X**	**All**
BSM	0.988	0.466	-0.327	-0.930	-1.709	0.022
HG	0.391	-0.274	-0.357	-0.381	-0.391	-0.104
P-BSM	-0.291	-0.197	-0.237	-0.277	-0.463	-0.278
P-HG1	0.187	-0.043	-0.057	-0.155	-0.416	-0.029
P-HG2	0.240	-0.060	-0.063	-0.015	0.070	0.058
P-HG3	-0.028	-0.026	0.024	-0.093	-0.219	-0.046
SP-BSM	-0.243	-0.276	-0.303	-0.284	-0.460	-0.294
SP-HG1	-0.201	0.135	0.346	0.265	-0.502	0.017
SP-HG2	0.126	0.087	0.025	-0.227	-0.483	-0.021
SP-HG3	-0.155	-0.335	-0.255	-0.249	-0.352	-0.249

Table VII: Monthly RMSE

Monthly RMSE by moneyness and by maturity in percentage of BSM model's RMSE. BSM is the Black-Scholes Model, HG is the Homoscedastic Gamma Model, P-BSM and P-HG are practitioner's versions of these models where volatility is quadratic in moneyness. P-HG1 is a version where the linear term is zero (i.e. $\gamma_1 = 0$), P-HG2 imposes that β is the risk-neutral skewness (see text) and P-HG3 is unrestricted. SP-BSM and SP-HG are smoothed version of these models where the shape of the quadratic IV curve is constant through the sample.

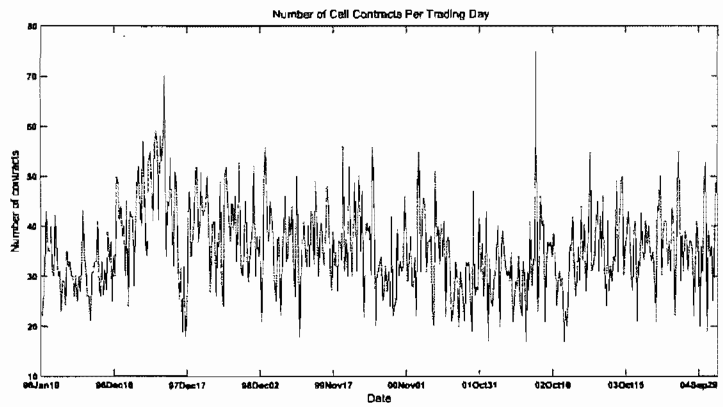
(a) Monthly In-Sample RMSE

Model	**Moneyness**					
	S/X<0.95	**0.95<S/X<0.975**	**0.975<S/X<1**	**1<S/X<1.025**	**1.025<S/X**	**All**
HGM	0.737	0.894	0.924	0.774	0.606	0.757
P-BSM	0.615	0.897	0.914	0.713	0.463	0.674
P-HG1	0.628	0.832	0.834	0.728	0.662	0.700
P-HG2	0.631	0.834	0.868	0.713	0.544	0.679
P-HG3	0.540	0.777	0.796	0.661	0.465	0.605
SP-BSM	0.647	0.891	0.919	0.750	0.513	0.700
SP-HG1	0.676	0.862	0.867	0.753	0.740	0.747
SP-HG2	0.641	0.835	0.877	0.752	0.603	0.701
SP-HG3	0.571	0.800	0.808	0.686	0.506	0.632

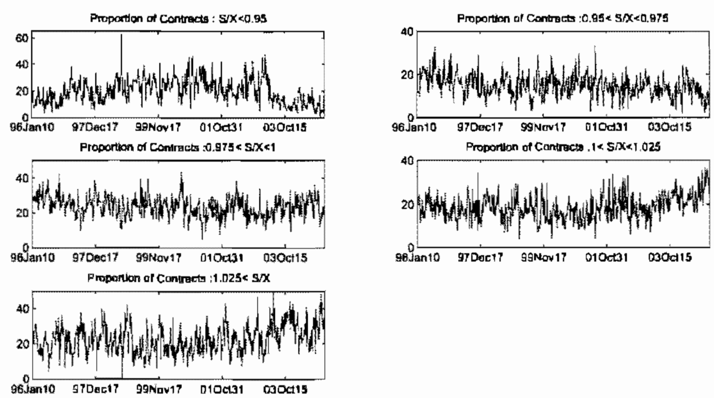
(b) Monthly Out-of-Sample RMSE

Model	**Moneyness**					
	S/X<0.95	**0.95<S/X<0.975**	**0.975<S/X<1**	**1<S/X<1.025**	**1.025<S/X**	**All**
HG	0.908	1.019	0.996	0.902	0.746	0.913
Q-BSM	0.855	0.988	0.991	0.925	0.761	0.891
Q-HG1	0.880	1.028	0.991	0.904	0.776	0.906
Q-HG2	0.867	1.020	0.996	0.902	0.752	0.897
Q-HG3	0.841	1.001	0.990	0.918	0.764	0.886
SQ-BSM	0.852	0.990	0.990	0.933	0.769	0.892
SQ-HG1	0.898	1.024	0.986	0.903	0.822	0.919
SQ-HG2	0.861	0.977	0.968	0.907	0.789	0.889
SQ-HG3	0.833	0.998	0.979	0.913	0.770	0.881

Figure 3.1: Number of call option contracts at each date

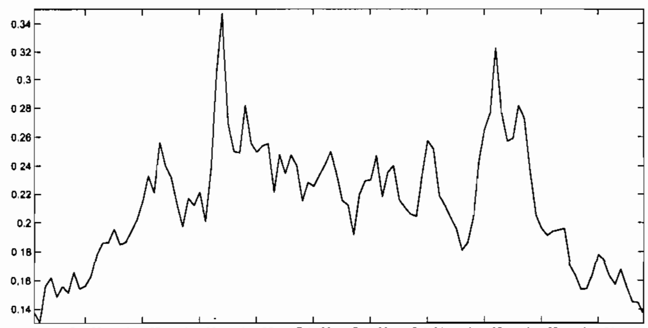


(a) Total number of contracts.

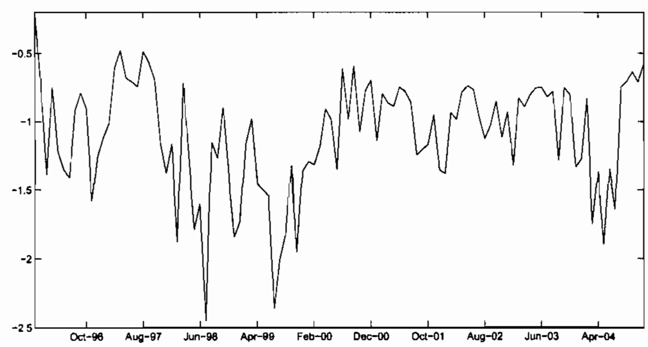


(b) Proportion of contracts in each moneyness category.

Figure 3.2: Time series of implied volatility and implied skewness from the smoothed version of the SP-HG2 model. This is a practitioner's version of the Homoscedastic Gamma model where the IV curve is restricted to depends only on the (constant) excess kurtosis.

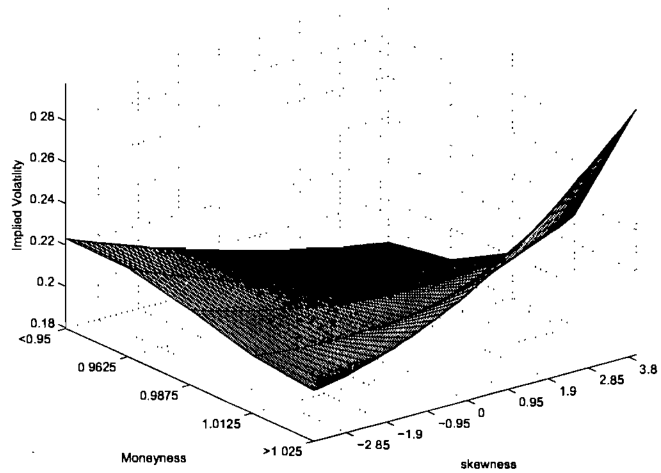


(a) Implied Volatility

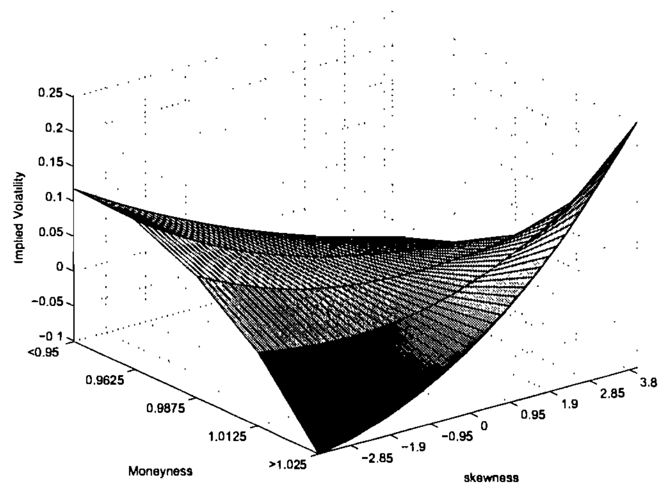


(b) Implied Skewness

Figure 3.3: Implied Volatility curves across values of skewness in level (Panel (a)) and in percentage deviation relative to the benchmark (i.e. zero skewness) BSM case (Panel (b)) , The grid covers 41 equidistant values of skewness and moneyness is defined as $\frac{\ln(S/K)(-r\tau)}{\bar{\sigma}\sqrt{\tau}}$ to correct for maturity differences.

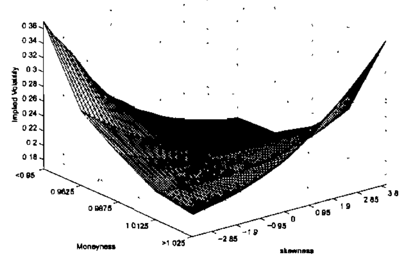


(a) Level

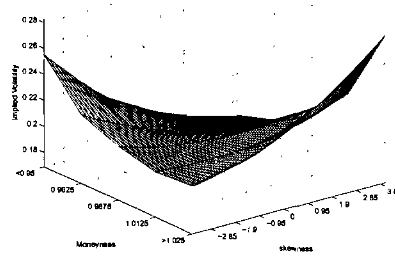


(b) Percentage Deviation from BSM

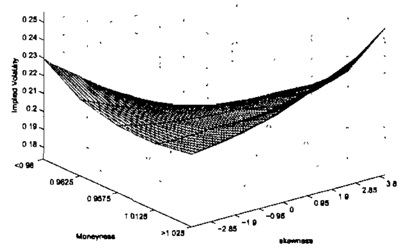
Figure 3.4: Implied volatility and skewness surfaces for different maturity categories where moneyness is defined as $\ln(S/K)(-r\tau)$. Maturity groups are defined using settlement dates.



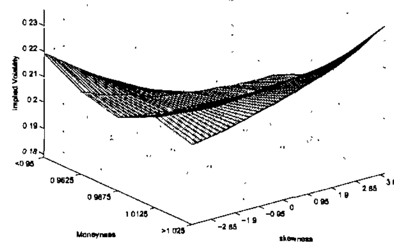
(a) Month 1



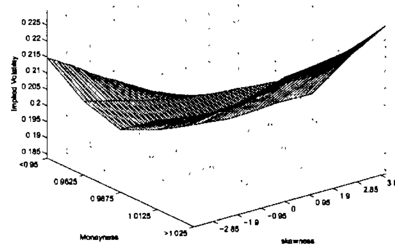
(b) Month 2



(c) Month 3

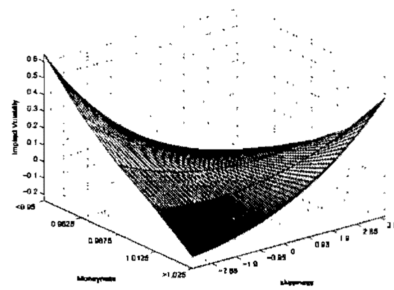


(d) Months 4 to 6

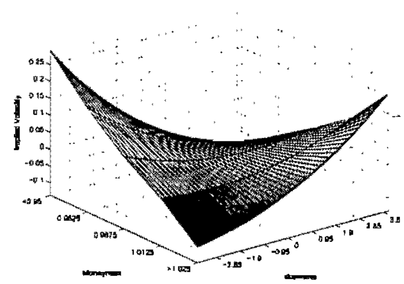


(e) Months 7 to 9

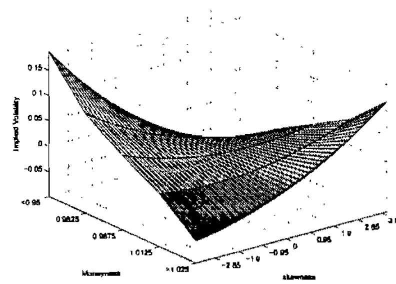
Figure 3.5: Deviations of implied volatility and skewness surfaces from the BSM IV values for different maturity categories. Moneyness is defined as $\ln(S/K)(-r\tau)$ and maturity groups are defined using settlement dates.



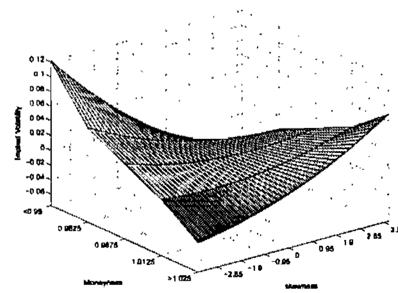
(a) Month 1



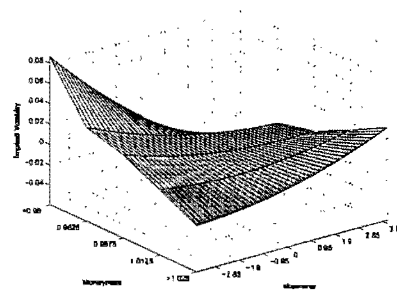
(b) Month 2



(c) Month 3

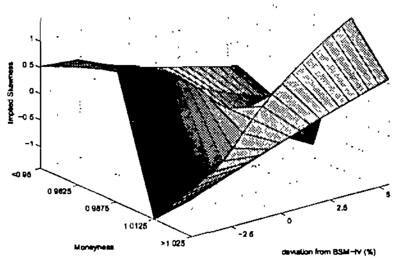


(d) Months 4 to 6

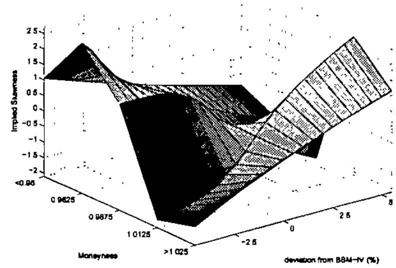


(e) Months 7 to 9

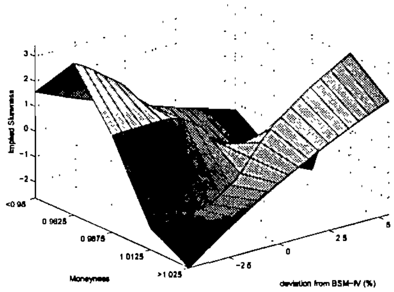
Figure 3.6: Implied skewness curve for different values of volatility, in percentage deviation from BSM IV values, for different maturity groups. Moneyness is defined as $\ln(S/K)(-r\tau)$ and maturity groups are defined using settlement dates.



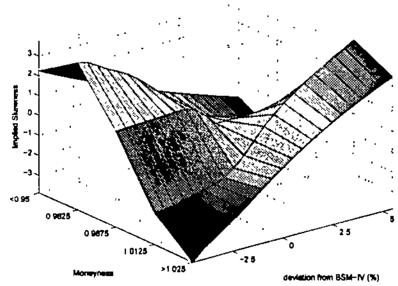
(a) Month 1



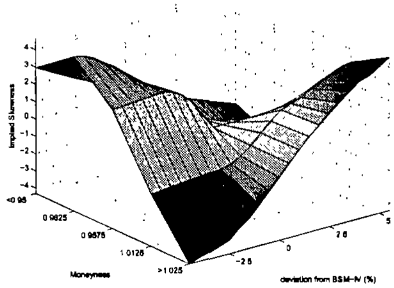
(b) Month 2



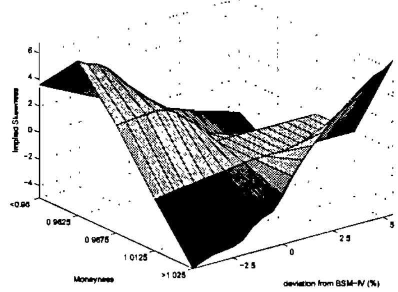
(c) Month 3



(d) Months 4 to 6

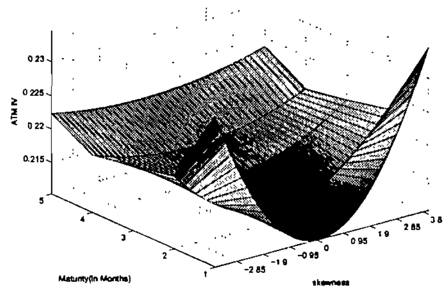


(e) Months 7 to 9

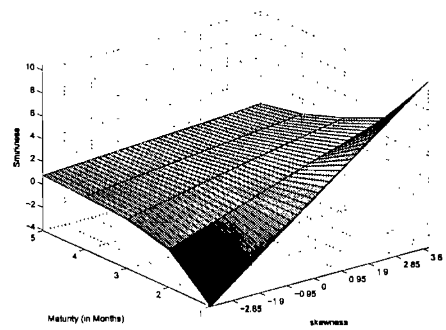


(f) Months 10 to 12

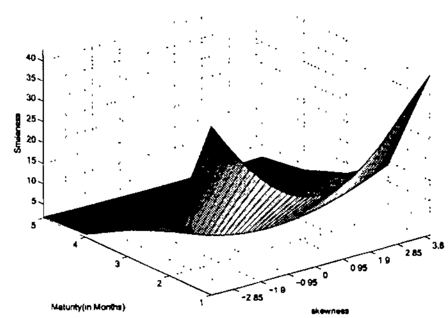
Figure 3.7: Time-series average of estimates of $\Theta_t = (\sigma_t, \gamma_{1,t}, \gamma_{2,t})$ from the P-HG3 (unrestricted) model but for different values of skewness. The parameters govern the IV curve: $\sigma_{i,t} = \sigma_{I0,t}(1 + \gamma_{1,t}\xi_i + \gamma_{2,t}\xi_{i,t}^2$.



(a) $\sigma_{I0,t}$ and Skewness

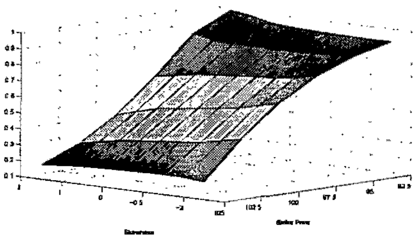


(b) $\gamma_{1,t}$ and Skewness

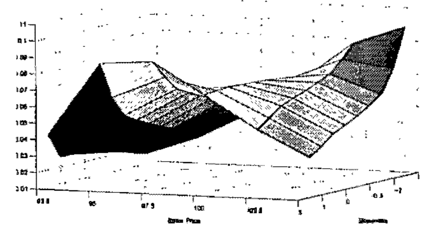


(c) $\gamma_{2,t}$ and Skewness

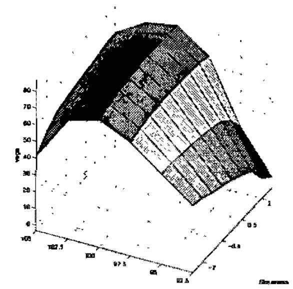
Figure 3.8: Option prices sensitivities.



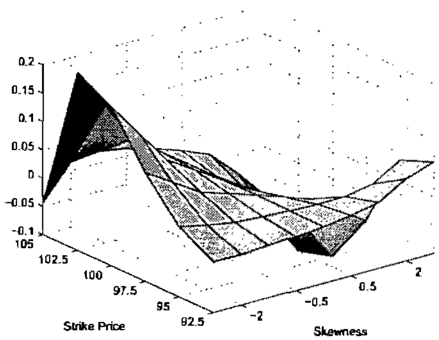
(a) First derivative with respect to stock price, in level.



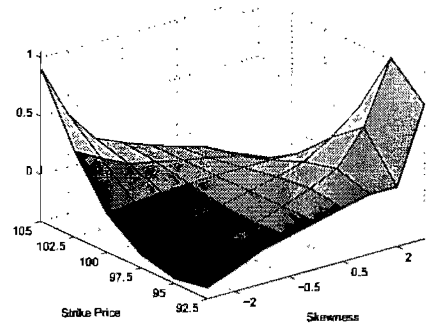
(b) Second derivative with respect to stock price, in level.



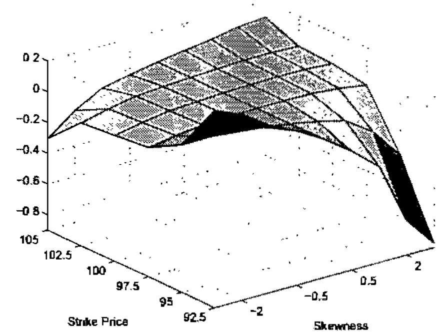
(c) Derivative with respect to volatility, in level.



(d) First derivative with respect to stock price, in percentage deviation.

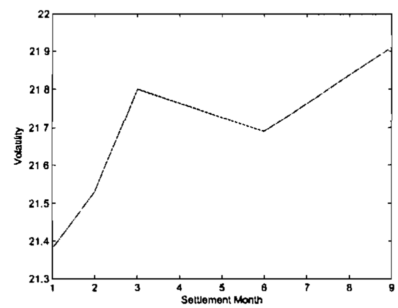


(e) Second derivative with respect to stock price, in percentage deviation.

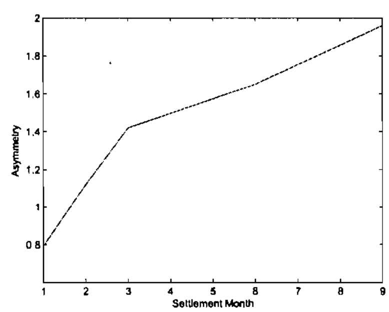


(f) Derivative with respect to volatility , in percentage deviation.

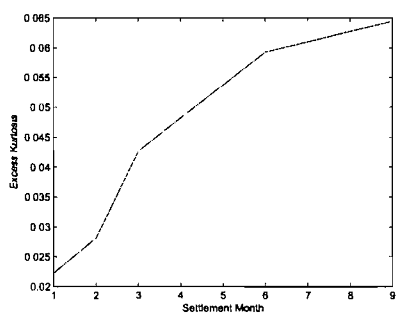
Figure 3.9: Term Structure of implied volatility, (minus) the implied skewness and (minus) the implied excess kurtosis from the SP-HG2 model.



(a) Implied Volatility



(b) Implied Skewness



(c) Implied Kurtosis

Discussion Générale Et Conclusion

La liquidité des marchés financiers est le thème central de cette thèse, bien qu'il soit abordé de manières bien différentes dans chaque article. Dans le premier chapitre je m'intéresse directement à l'importance des contraintes de crédit auxquelles font face les intermédiaires financiers. Dans ce context la source d'un manque de liquidité ou d'un risque d'illiquidité est bien différente des sources généralement admises. Tout d'abord, il ne s'agit pas d'information asymétrique. C'est-à-dire qu'il n'est pas nécessaire que certains agents ou intermédiaires n'aient qu'une connaissance imparfaite au sujet de l'actif financier qui nous intéresse. Il ne suffit pas non plus que les coûts d'inventaires des intermédiaires soient variables et risqués. C'est-à-dire les coûts associés à la possibilité qu'un intermédiaire ne puisse pas écouler sur le marché tous les actifs acquis alors qu'il offrait des services de liquidité à ses clients. La littérature théorique récente met plutôt l'emphase sur les limites à l'emprunt et le risque de défaut que doivent confronter les intermédiaires. Cependant, il était jusqu'à maintenant loin d'être clair comment obtenir une mesure des conditions d'emprunt et, subséquemment, de mesurer leurs impacts sur le prix des actifs financiers. C'est à ce niveau que nous apportons une première contribution. Intuitivement, nous identifions sur le marché des obligations du Trésor américain des paires d'obligations qui ne diffèrent que par l'aise avec laquelle les intermédiaires peuvent les financer. Ainsi, les variations de la différence entre les prix de ces deux obligations peuvent être attribuées aux variations dans les conditions d'emprunt des intermédiaires. Alors que ces variations pourraient n'avoir qu'un impact limité sur l'économie, notre seconde contribution est de montrer leur importance économique. En particulier, nous établissons un lien avec des variations substantielles de la prime de risque, et donc les prix, d'un éventail d'actifs financiers. En conclusion, alors que les primes d'illiquidité sont généralement admises comme spécifiques à