

Modeling Sovereign Yield Spreads: A Case Study of Russian Debt

Darrell Duffie
Lasse Heje Pedersen
Kenneth J. Singleton*

First Version: August 30, 1999, Current Version: September 24, 2001

ABSTRACT

We construct a model for pricing sovereign debt that accounts for the risks of both default and restructuring, and allows for compensation for illiquidity. Using a new and relatively efficient method, we estimate the model using Russian dollar-denominated bonds. We consider the determinants of the Russian yield spread, the yield differential across different Russian bonds, and the implications for market integration, relative liquidity, relative expected recovery rates, and implied expectations of different default scenarios.

JEL classification: C5, E4, F34, G12.

*Graduate School of Business, Stanford University, Stanford, CA 94305-5015. Emails: duffie@stanford.edu, lasse@stanford.edu, and ken@future.stanford.edu. We thank Martin Jacobsen, Andrei Khinchuk, David Lando, Vladimir Semyonov, and Len Umantsev for discussions, as well as seminar participants at the Anderson School at UCLA, National Bureau of Economic Research, Princeton University, Stanford University, and Yale University. Pedersen gratefully acknowledges financial support from the Danish Research Academy.

This paper develops a model of the term structure of credit spreads on sovereign bonds that accommodates: *(i)* Default or repudiation: The sovereign announces that it will stop making payments on its debt; *(ii)* Restructuring or renegotiation: The sovereign and the lenders “agree” to reduce (or postpone) the remaining payments; and *(iii)* A “regime switch,” such as a change of government or the default of another sovereign bond that changes the perceived risk of future defaults. We build on the framework of Duffie and Singleton (1999), showing that cash flows promised by a sovereign bond can be discounted using a default-adjusted short-term discount rate that reflects the mean arrival rate, and associated losses in market value upon arrival, of each of the aforementioned types of credit events. Since a sovereign credit event is often not a “liquidation event,” the model allows for continued trading (and pricing) of a sovereign instrument through credit events, possibly after write-downs in face value or cash distributions to creditors on credit event dates.

Additionally, we accommodate the possibility that bonds issued by the same sovereign, of exactly the same type but possibly of different maturities, may be priced in the market using different discount factors. Reasons for this may include: *(i)* Bond covenants may not include cross-default clauses that would force, upon the default of one bond, the simultaneous default of other bonds of the same type, but of a different maturity. For various strategic reasons related to internal or external political or economic considerations, sovereign issuers may choose to default on, or to renegotiate, the terms of one bond (or one set of bonds), but not on others. *(ii)* Some sovereigns (such as Turkey and the Philippines) have issued Eurobonds governed by different legal systems (those of the U.K. and U.S.) that have different collective-action clauses and, hence, differ with respect to the ease of creditor-initiated restructurings (Eichengreen and Mody (2000)). *(iii)* Portions of the outstanding debt may have been issued during different political regimes. In this situation, the current regime may feel, or be perceived to feel, a stronger obligation to make contractual payments on debt issued during its own regime. *(iv)* Finally, for these reasons and because of possible clientele trading patterns, high transaction costs, or asymmetrically informed traders (including government insiders), different bonds may have different illiquidity.

As an illustration of our valuation framework, we estimate a model of the term structure of credit spreads for bonds of the Russian Ministry of Finance (MinFin) over a sample period encompassing the default on do-

mestic Russian GKO bonds in August, 1998.¹ For this purpose, our model is parameterized with the composite credit-and-liquidity spread of one of the Russian MinFins, serving as a “benchmark.” All other MinFins are allowed to have idiosyncratic components to their discount rates relative to this benchmark. Importantly, these idiosyncratic factors enter directly into the discount rates, rather than being treated as additive pricing or measurement errors, as has become common practice in the literature on dynamic term structure models.² Initially, our model is estimated using data up to, but not including the 1998 Russian default. Using the parameters and implied spreads, we examine a variety of implications of the estimated model, including the determinants of the spread, the degree of integration of the markets for different Russian sovereign bonds, differences in both the liquidities and expected default recoveries of different bonds, and the implied expectations of investors about different default scenarios. We also study the correlations of these spreads with various macro-economic time series including the reserves of the Russian Central Bank and the price of oil.

Identifying an econometric model of spreads with a sample period that includes a major sovereign credit event is challenging, because such events are rare — there is only one in our Russian sample! Moreover, as outlined above, there may be effective reductions of principal due to the credit event that may only be revealed at later dates, possibly as late as maturity. Our approach to modeling “through a credit event” in Russia is decidedly pragmatic. We extend our basic pricing framework, under admittedly strong independence assumptions, to allow for the possibility that creditors do not know with certainty the implications of a credit event for the eventual write-down of the bond. Instead, they formulate prices based on their expected write-down. Then, through the pricing model, we are able to back out time series of both a default-adjusted credit spread and the creditors’ time-varying expectations about their eventual payoff on each MinFin bond. Our time-series analysis suggests that the GKO default lead to expectations of substantial write-downs on MinFins, but that investors subsequently grew increasingly optimistic that there would be (nearly) full recoveries of face value. Moreover, markets seemed to “price in” the political risks related to the Russian government’s expressed views about the relative strengths of their commitments to redeem the MinFins for face value at maturity.

In estimating our pricing models, we use a novel efficient estimation methodology, based on a new approximation to the likelihood function³ for a large class of affine diffusions. The basic idea of our approach is to divide the

state vector X_t into the sub-vector X_t^I that drives the stochastic volatility of X_t and the sub-vector X_t^D of remaining state variables, as in the canonical representation of affine diffusion models proposed by Dai and Singleton (2000). For the special case in which the elements of X_t^I are independent square-root processes,⁴ the conditional density function of X_t^I is known to be the product of non-central chi-square distributions (Cox, Ingersoll, and Ross (1985)). Furthermore, the conditional distribution of X_t^D , given both X_{t-1} and the intervening sample path $\{X_s^I, t-1 < s \leq t\}$, is normal with conditional moments that are known in closed form. These properties allow us to derive an approximation to the density of X_t conditional on X_{t-1} that leads to computationally tractable estimators for the models that we have examined.

The remainder of this paper is organized as follows. Section I discusses the nature of credit risk in sovereign debt markets, and then presents our model for pricing sovereign debt. Section II introduces the parametric special case underlying our analysis of Russian bonds and presents our estimation methodology. Section III presents our empirical results for Russian bond data.

I. Pricing Sovereign Debt

The nature of the economic factors that drive default decisions, and the range of alternative restructuring and recovery plans considered in the event of default, typically differ between sovereign and corporate debt. To motivate our modeling strategy, we begin by discussing the nature of sovereign credit risk and approaches that have been taken to modeling this risk in the existing literature.

A. The Nature of Sovereign Credit Risk

A holder of sovereign debt may not have recourse to a bankruptcy code in the event of default. Moreover, a sovereign default is largely a political decision. Governments trade off the cost of making debt payments against reputation costs (Eaton and Gersovitz 1981), the costs of having assets abroad seized, and the costs of having international trade impeded (Bulow and Rogoff (1989b) and Gibson and Sundaresan (1999)). A sovereign rarely makes an outright default. Rather, it may force a restructuring or renegotiation

of its debt. Indeed, the same bond may be repeatedly renegotiated (Bulow and Rogoff 1989a). A sovereign also trades off the costs of default (or forced restructuring) of internal versus external debt. As we shall see, this trade-off has interesting implications for the pricing of different classes of Russian debt. We are not aware of any theoretical papers that address this issue. Indeed, most papers on sovereign debt address only external debt. Kremer and Mehta (2000) study a model in which a government is more inclined to default if a large proportion of its debt is held by foreigners.

Though, for these reasons and others, sovereign debt is notably different than corporate debt, these instruments have typically been priced using similar structural or reduced-form models of a single “default” credit event. In the former, default occurs when incentives suggest that it is optimal for the issuer to default, or when payment on debt is impossible. Default is sometimes modeled as the event that, at maturity, there are insufficient assets to pay down the debt, as in Merton (1974), or the event that the debtor’s cash-flows or asset-liability ratio falls below some cut-off level for the first time, as in Fischer, Heinkel, and Zechner (1989). In reduced-form models, the default time has an exogenously specified intensity process, and therefore arrives as a “surprise” (at an inaccessible stopping time) that is not endogenously linked to (observed) decision variables of the debtor (Jarrow and Turnbull (1995), Lando (1998), Duffie and Singleton (1999)). The exogenous intensity process may, however, be allowed to depend on observable issuer-specific or macro-economic variables.

Structural models, which directly capture the default incentives and solvency of the issuer, can be problematic when empirically modeling sovereign debt. For instance, Deutsche Bank Research (1998b) states that “the Russian budget itself is of course largely fictitious in the sense that it can make the deficit pretty much what you want.” Beyond this measurement issue, the incentives of a sovereign to default are relatively complex. When a sovereign defaults, it may lose assets held abroad, but assets held within the country need not be seized as collateral. Rather, the country loses reputation for its failure to make payments, which worsens its access to international capital markets, and may impede international trade. Politicians trade off these costs against the cost of making payments on the debt, along with other political considerations (including their own personal incentives). Deutsche Bank Research (1998a) writes: “We continue to maintain that a default depends far more on Russia’s willingness to pay versus its ability to pay its debt.” The incentives of a corporation are usually simpler: It defaults (in

theory) when it cannot make its payments, or when equity holders find that the market value of their shares has dropped to zero. Gibson and Sundaresan (1999) present a formal model in which there are significantly different optimal default strategies for sovereign and corporate borrowers, in part because sovereign debt is not directly covered by a bankruptcy code.

In the light of these and other pragmatic considerations, studies of sovereign debt pricing have focused primarily on models based on an exogenously specified intensity process.⁵ Merrick (1999) calibrates a discrete-time model (to Russian and Argentinian bonds) that can be reinterpreted as a model with a constant (state-independent) intensity. More generally, Keswani (1999) and Pagès (2000) apply special cases of the modeling framework of Duffie and Singleton (1999) to data on Latin American Brady bonds, and Dullmann and Windfuhr (2000) apply a similar framework to pricing European government credit spreads under the EMU. Implicit in these formulations are the assumptions that holders of sovereign debt face a single credit event — default, with liquidation upon default — and that the bonds issued by a given sovereign are homogeneous with regard to their credit characteristics.

The case of Russian bonds, and in particular the Russian MinFins on which we focus in our empirical analysis, highlights the potential limitations of applying corporate-bond pricing models to sovereign debt. Table I lists a selection of the Russian U.S.-dollar-denominated debt that is studied in Section III. In 1993, five dollar-denominated MinFin bonds were issued as payment to Russian exporters for Vnesheconombank accounts that were frozen in 1991. In 1996, two additional MinFins were issued, and Russia issued its first Eurobond. Several more Eurobond issues followed in 1997 and 1998.

Based on past experiences with Soviet-era debt and the debts of other sovereigns, investors in these bonds would surely have recognized that outright “default” was not the only possible, or indeed even the most likely, credit event. More likely would be a restructuring, with possible additional restructurings prior to maturity. In fact, between April, 1993 and August, 1999, the “Paris Club” of Western government creditors restructured its Soviet-era debts with Russia five times. Moreover, as part of any restructuring, there is the possibility of write-downs of principal. The Russian Federation defaulted on the principal payments at maturity of the MinFin 3 bonds that matured in May, 1999 (essentially, a write-down at maturity). More recently, the restructuring of debt held by the Russian “London Club” of Western commercial lenders announced in early 2000 called for both write-downs

Issue	Issue Date	Maturity	Coupon	Amount Issued (Billions of US\$)
MinFin 3	5/14/1993	5/14/1999	3.00	1.32
MinFin 4	5/14/1993	5/14/2003	3.00	3.38
MinFin 5	5/14/1993	5/14/2008	3.00	2.84
MinFin 6	5/14/1996	5/14/2006	3.00	1.75
MinFin 7	5/14/1996	5/14/2011	3.00	1.75
Eurobond-01	11/27/1996	11/27/2001	9.25	2.40
Eurobond-07	6/26/1997	6/26/2007	10.00	1.00

Table I: Contractual Characteristics of Russian Dollar-denominated MinFins and Eurobonds.

and obligor status. In 1997, the London Club restructured its Soviet-era debt into two securities: \$6 billion of principal notes (Prins) and \$20 billion of interest-arrears notes (IANs). The subsequent restructuring announced in 2000 called for reductions in principal of 37.5% and 33% of the Prins and IANs, respectively, and a change in obligor status from Vnesheconombank to the Russian Federation through the issue of Eurobonds (IMF (2000a)).

Additionally, the legal covenants of these bonds and the characteristics of the markets in which they were traded presented investors with a rather complex menu of credit and liquidity risks. Eurobonds are much like U.S. corporate bonds in that they have cross-default triggers. That is, failure to make payments on one bond allows the owners of other bonds to accelerate their bonds, requiring immediate repayment of the principal. If a company defaults on one bond, then other creditors have an incentive to join the restructuring, for otherwise the restructuring agreement between the equity holders and holders of bonds in default might severely lower the value of the bonds that are not (yet) in default. Hence, for corporate bonds, cross-default triggers protect bondholders from discrimination. For sovereign bonds, on the other hand, cross-default triggers may have little effect, because the owners of other bonds may have little incentive to exercise their right to accelerate their bonds and join in a restructuring. Even among Eurobonds, the nature of this “right to accelerate” differs. Russian Eurobonds are governed by U.K. laws which, relative to the U.S. system, have collective-action clauses that make it easier for bondholders to initiate restructurings.

In contrast, although MinFins are denominated in U.S. dollars, they are

technically domestic debt under the jurisdiction of Russian law and have no cross-default clauses. This gives the Ministry of Finance some discretion in deciding which creditors are most severely affected by a credit event (for example, which bonds will default), as well as discretion over the nature of a credit event, including the effective recovery rate.

Another feature of Russian MinFins is the importance of “political” risk. The Ministry of Finance has drawn a clear distinction between their commitments to fulfill the principal obligations of “Soviet-era” versus “post-Soviet-era” debt. MinFins 3 through 5, as indicated in Table 1, were issued in 1993, and hence are technically “post-Soviet-era” debt, but were issued to address “Soviet-era” problems. There has been no explicit commitment by Russia to repay their principals at maturity. In fact, Russia defaulted on the principal obligation of the MinFin 3 bonds in 1999. In contrast, the Russian Ministry of Finance has publicly committed to fulfill all of its payment obligations on the MinFin 6 and MinFin 7, as well as on the Eurobonds.

The effects of the Russian default on GKO in August, 1998 on the prices of MinFins are displayed in Figure 1, which shows the prices of five MinFin bonds from 1/2/1998 to 11/17/2000, relative to their prices on 7/31/1998 (just prior to default). Note, first of all, that MinFin prices fell by approximately 80% during the week of the Russian default, suggesting that market participants were indeed surprised by this credit event right up to the time of the event itself. This is roughly consistent with default arrival at an inaccessible stopping time, as in our intensity-based modeling approach.

Second, the general patterns of relative prices around the time of default suggest that a fractional recovery of market value is not an unreasonable assumption relative, say, to a typical alternative assumption that bondholders would recover an exogenous fraction of the bond’s face value. Third, the fact that the price of the MinFin 3 did not approach its face value just prior to maturity suggests that many investors believed, prior to May, 1999, that they would not receive 100% of face value at maturity of these MinFins. Essentially, this bond had already gone into default in August, 1998, although the precise terms of default were not known at that time. Likewise, we could think of the other MinFins as being in default after August, 1998. These bonds have recently recovered substantially, however, which we interpret as a sign of investors being more optimistic about the impact of the (implicit) default of August, 1998 on the eventual payments at maturity. This optimism was also reflected in the index of Russian stock prices shown in Figure 1.

Fourth, the MinFins recovered about 20–30% of their previous market



Figure 1. Prices of Russian MinFin 3–7 bonds over the time period from 1/2/1998 to 11/17/2000, normalized so that they all have a market value of 100 at 7/31/1998 (just prior to default). Source: Datastream.

values after the credit event of August, 1998, whereas the Eurobonds (not shown) recovered around 30–40% of their market values. Additionally, Figure 1 shows that the MinFin 6 and MinFin 7 bonds have recovered larger fractions of their pre-default values than have the MinFin 4 and 5 bonds. Hence, since the August, 1998 default, investors may have been viewing Eurobonds and post-Soviet MinFins as being of higher quality than Soviet-era MinFins. We shall later examine whether this viewpoint was reflected in pre-default prices.

B. A Model for Pricing Sovereign Debt

Initially, we characterize the price, P_t , at any time t before default, of a sovereign security that promises to pay a single, possibly random, amount Z at some time $T > t$. We fix a probability space and an increasing family $\{\mathcal{F}_t : t \geq 0\}$ of information sets defining the resolution of information over time, with $Z \in \mathcal{F}_T$. We take as given a short-rate⁶ process r and an equivalent martingale measure Q . (See Harrison and Kreps (1979) and Delbaen and Schachermayer (1999) for the essential equivalence of the existence of such a measure and the absence of arbitrage.) That is, a security paying some amount \tilde{Z} at some time (possibly a stopping time) S has a market value at any time t before S of $E_t^Q \left[e^{\int_t^S -r(u) du} \tilde{Z} \right]$, where E_t^Q denotes \mathcal{F}_t -conditional expectation with respect to Q . Uniqueness of the risk-neutral measure Q is not to be expected in our setting.

For the case of a single credit event (default), Duffie and Singleton (1999) showed that for pricing purposes we can treat the cashflow X as default-free and allow for default risk by replacing the discount rate $r(t)$ with the “default-adjusted” discount rate $r(t) + h(t)L(t)$, where h is the risk-neutral intensity of default and L is risk-neutral expected fractional loss in market value in the event of default. In modeling sovereign spreads, we extend this model to one that accommodates many types of credit and liquidity events, in particular the types (i)–(iii) of events enumerated in the introduction. This yields a pricing formula similar to that derived (in a less general setting) by Schönbucher (1998), who considers multiple defaults.

To accommodate multiple types of credit events, we let N_t be the number of credit events of any type that have happened by time t , where N is a counting process with risk-neutral intensity process⁷ h . We assume that, at the n -th credit event, the promised payment of the security is lowered to a fraction, Y_n , of its pre-credit-event value. For example, a sovereign issuer may simply unilaterally announce at the first credit event that the principal that it recognizes as an obligation is reduced by 50%, in which case $Y_1 = 0.5$. This leaves open the possibility of a further repudiation at a later date. Also, at the n -th credit event, investors may receive, or prospectively value, cash flows paid by the issuer in lieu of the repudiated portion of the debt, with a current market value that is a fraction W_n of the pre-credit-event market value of the obligation.⁸ We assume that h , $\{Y_n\}$, and $\{W_n\}$ are exogenously given. The total effective fractional loss in market value caused by the n -th credit event is thus $1 - Y_n - W_n$. At time t , the risk-neutral expected fractional

loss associated with the next credit event, were it to occur immediately, is denoted L_t .⁹ Hence, the process describing the risk-neutral mean fractional loss rate (due to credit events) is hL .

We also allow liquidity to affect pricing. As in Duffie and Singleton (1999), we make the simplistic assumption that illiquidity of the security translates into a fractional cost rate of l , where l is a predictable process. Hence, the total mean loss rate of the security due to credit events and illiquidity is $hL + l$. If the repudiated fraction Y_n is observable at the event time τ_n (that is, $Y_n \in \mathcal{F}_{\tau_n}$), then, at time t the security is worth the fraction $\prod_{n \leq N(t)} Y_n$ of an otherwise identical security that has not yet experienced credit events. This gives the intuition for the following pricing formula.

Theorem 1 *Let $R = r + s$, where*

$$s = hL + l. \tag{1}$$

Suppose that $V_t = E_t^Q \left(e^{-\int_t^T R_u du} Z \right)$ is well defined, and that (almost surely), the process V does not jump at a credit event time.¹⁰ If the repudiated fraction Y_n is observable at the event time τ_n , then,

$$P_t = E_t^Q \left(e^{-\int_t^T R_u du} Z \right) \prod_{n \leq N_t} Y_n. \tag{2}$$

We refer to the process s in (1) as the short (or instantaneous) spread. The condition that V does not jump at a credit event time is automatically satisfied in our model of Russian bond prices because V_t depends on the current level of a diffusion process.

Our treatment counts all credit events with a single process N . One could also distinguish among different types of credit events. Suppose, for instance, that, under the given risk-neutral measure Q , liquidation events have intensity process h^1 and expected fractional loss process L^1 , that restructuring likewise has intensity h^2 and expected fractional loss L^2 , and that regime switches occur with intensity h^3 and expected fractional loss L^3 . Assuming that not more than one of these events can happen at any one time, the total risk-neutral intensity process of a credit event is $h = h^1 + h^2 + h^3$, and the expected fractional loss at such an event is the intensity-weighted average, $L = \frac{h^1}{h}L^1 + \frac{h^2}{h}L^2 + \frac{h^3}{h}L^3$. Hence, by Theorem 1, the short spread is

$$s = h^1 L^1 + h^2 L^2 + h^3 L^3 + l. \tag{3}$$

Ideally, sufficient data would be available to identify each term of (3) separately, but this is not the case for our analysis of Russia and so we shall not attempt such a disentangling of credit event types in our empirical analysis.

Even without formal econometric analysis, the behavior of the prices of MinFin 3 shown in Figure 1 suggests that allowing for the risk of write-downs that are instigated by the default on GKO is critical for pricing Russian bonds. As the MinFin 3 matured in 1999, its price approached about a third of its face value. In a purely intensity-based model of credit risk, however, because MinFin 3 was not in default, its price must approach face value at maturity. Not only does the observed pattern call for a write-down, but it challenges the strong assumption of Theorem 1, that the fraction Y_n of the notional that is recovered due to the n -th credit event is observable at the associated credit event time τ_n . An empirically more plausible view is that, as a consequence of the default on GKO in August, 1998, bond prices reflected investors' expected payment at maturity, and that investors' expectations were subsequently revised upwards over time toward one-third of face value as market and economic conditions changed.

To formalize the idea of "expected payments at maturity," we allow for bond investors to learn more about the extent of the write-down after the time of the credit event itself, including the possibility that the final resolution of the impact of a re-structuring is learned only at maturity. Although such learning substantially increases the difficulty of pricing sovereign bonds, we can extend our pricing relation under the (unrealistically strong) assumption of *Independent Recovery*: Under Q , $\{Y_1, Y_2, \dots\}$ are independent, and independent also of $\{Z, r, l, N, W_1, W_2, \dots\}$. With this assumption, our pricing result extends as follows.

Theorem 2 *Let $R = r + s$, where $s = hL + l$. Suppose that $V_t = E_t^Q \left(e^{-\int_t^T R_u du} Z \right)$ is well defined, and that (almost surely), V does not jump at a credit event time. Under the assumption of *Independent Recovery*,*

$$P_t = E_t^Q \left(e^{-\int_t^T R_u du} Z \right) \prod_{n \leq N_t} E_t^Q(Y_n). \quad (4)$$

The *Independent Recovery* assumption is strong because, for instance, as economic conditions in Russia improve, we would expect both that $E_t^Q(Y_n)$ is revised upward and that the credit-event arrival intensity h declines. The independence assumption does, however, allow for state-dependent write-

downs $E_t^Q(Y_n)$ that substantially increase the flexibility of our model in capturing market price movements. In particular, subsequent to the first credit event, when at least one $E_t^Q(Y_n)$ term is present in (4), bond prices may jump up or down as investors revise their expectations about the consequences of past credit events for eventual receipts at maturity. Such jumps might occur at any time, including just prior to the maturity of the bond. A notable example of the type of jumpiness that may capture with stochastic $E_t^Q(Y_n)$ is the rather saw-toothed pattern of the price of MinFin 3 in 1999 prior to its default at maturity (Figure 1). We note that jumps in $E_t^Q(Y_n)$ are conceptually distinct from jumps in prices that might occur as a consequence of jumps in the short spread s , due say to a jump in the risk-neutral intensity of credit events. Our conceptual framework extends immediately to allow for such jumps in R , using results of Duffie, Pan, and Singleton (2000).

These theorems give the price of a security with a single promised cash flow. The price of a security with multiple cash flows (such as coupons and principal) is, under our assumptions of exogenous intensity, loss rate, and liquidity, just the sum of the prices of the individual cash flows. Further, these pricing relations allow for different write-down factors, Y^c and Y^f , for coupons and principal, respectively. For our empirical analysis we set $Y^c = Y^f$, thereby focusing on a partial or full repudiation of the principal of a coupon bond, with the original terms otherwise unchanged. If, for this case, the fractional market values W^c and W^f paid in lieu of the repudiated portions of the coupons and principal, respectively, are identical, then a single short-spread process, s can be used to price all of a bond's promised cash flows. Each distinct bond issue may be valued by its own distinct spread process, however, as in our subsequent parametric pricing model for Russian bonds.

An attractive feature of the pricing relations (2) and (4) is that valuation can proceed using familiar term-structure analytics once the functional relations between the spread, s , the expected repudiation $E_t^Q(Y_n)$ (in the case of Theorem 2), and the state vector are chosen. In particular, the tractability of affine term-structure models (Duffie and Kan (1996), Dai and Singleton (2000)) can be exploited for pricing sovereign debt. The next section presents one possible affine model of sovereign bond prices, the one that we will use for studying Russian bond data.

II. Specification and Estimation of a Parametric Pricing Model

Within the affine family of term-structure models we can, in principle, allow for general affine dependence among the state variables driving the riskless term structure and the sovereign short spread (bond by bond), subject to the requirement that the model gives well-defined prices. It is computationally convenient, however, to work with “recursive” models that allow estimation of the parameters governing the reference term structure in a first stage, followed by estimation of a model of the spread process, s , in a second stage. This section presents the recursive model that we use in our empirical analysis and introduces a new “approximate” maximum likelihood estimator that is applicable to a broad class of affine term-structure models, including the one that we will implement for Russian data.

A. Parametric Model of the Reference Curve

We use the term structure of U.S.-dollar LIBOR-quality swap yields as the reference curve for our analysis. Though swaps are defaultable, reliable pricing of their cash flows is obtained by using a standard default-free term-structure model, whose default-free short-rate process is replaced with a refreshed LIBOR-quality default-adjusted short rate r (Duffie and Singleton (1997), Duffie and Huang (1996), Collin-Dufresne and Solnik (2001)). Working with the treasury curve as the reference curve would give quantitatively similar results, given that the credit risk in Russian bonds is orders of magnitude larger than default risk in the swap market. Advantages of swap over Treasury yields are that they are unencumbered by the effects of repo specials and tax advantages, and are available on a constant-maturity basis.

We assume a two-factor affine model for the reference term structure, with the dynamics of the short rate, r , and the process, v , driving volatility given by:

$$d \begin{bmatrix} v_t \\ r_t \end{bmatrix} = \begin{bmatrix} K^{vv} & 0 \\ K^{rv} & K^{rr} \end{bmatrix} \left(\begin{bmatrix} \theta^v \\ \theta^r \end{bmatrix} - \begin{bmatrix} v_t \\ r_t \end{bmatrix} \right) dt + \sqrt{v_t} \begin{bmatrix} 1 & 0 \\ \Sigma^{rv} & \Sigma^{rr} \end{bmatrix} dW_t^{v,r}, \quad (5)$$

where $W^{v,r} = (W^v, W^r)^\top$ is a two-dimensional standard Brownian motion. This parameterization allows both for correlation of unconstrained sign between the factors (through both the drift and diffusion terms), and for

stochastic volatility in r induced by v . This is in contrast to the pure Gaussian or pure CIR-style two-factor models, which each accommodate one, but not both, of these features.

To complete the model, we specify the distribution of the risk-neutral¹¹ Brownian motion $\tilde{W}^{v,r}$ as:

$$dW_t^{v,r} = - \begin{bmatrix} \sqrt{v_t} & \\ & \sqrt{v_t} \end{bmatrix} \begin{bmatrix} \lambda^v \\ \lambda^r \end{bmatrix} dt + d\tilde{W}_t^{v,r}, \quad (6)$$

where (λ_v, λ_r) are the “market prices of risk.” Finally, we denote the parameter vector of the reference model by

$$\psi = (\theta^v, \theta^r, K^{vv}, K^{rv}, K^{rr}, \Sigma^{rr}, \Sigma^{rv}, \lambda^v, \lambda^r).$$

B. An Approximate Maximum Likelihood Estimator

From the literature on the estimation of affine term-structure models, we know the functional relation between the vector c_t of reference yields used in estimation and the state vector $X_t' \equiv (r_t, v_t)$ is $c_t = g(X_t; \psi)$, for a differentiable function g that is easily computed.¹² At a given parameter vector ψ , the model-implied state vector X_t^ψ is thus given by

$$X_t^\psi := g(\cdot; \psi)^{-1}(c_t), \quad (7)$$

assuming invertibility (which is not an issue in our application). Throughout, we use the notation “ $p(X | Z; \psi)$ ” to denote the conditional density of some random vector X given some random vector Z , at a parameter vector ψ (assuming this density exists); and likewise use “ $p(X; \psi)$ ” for the unconditional density. Letting $c = (c_1, \dots, c_T)$ denote the sequence of observed vectors of reference yields, standard change-of-variable arguments lead to the likelihood

$$p(c; \psi) = \prod_{t=1}^T p(X_t^\psi | X_{t-1}^\psi; \psi) \frac{1}{|\det Dg(X_t^\psi; \psi)|}, \quad (8)$$

where, for affine models, the Jacobian Dg can be easily computed.

Unfortunately, the transition density $p(X_t | X_{t-1}; \psi)$ of an affine diffusion X is not generally known in closed form, outside of special cases, such as Gaussian and independent square-root diffusions. We have found, however,

that a simple approximation to the likelihood function leads to reliable estimates in a large sub-class of affine models, including the models examined in this paper. Our approximation is based on the observation that, without loss of generality for a large family of affine diffusions, we can partition as $X = (X^I, X^D) \in \mathbb{R}_+^n \times \mathbb{R}^m$, where the distribution of X_t^I conditional on X_{t-1}^I is non-central chi-square and the distribution of X_t^D conditional on X_{t-1}^I and the entire path of X^I between dates $t-1$ and t is normal. Thus, to obtain an approximation to $p(X_t \mid X_{t-1}; \psi)$ we need only approximate the second of these distributions and, as discussed in Appendix B, this can be done reliably in a computationally tractable way. The resulting *ML* estimator is more efficient than the quasi-*ML* estimator (see, for example, Fisher and Gilles (1996) and Duffee and Stanton (2001)). Asymptotic efficiency of our *ML* estimator is achieved by the simulated-method-of-moments estimator proposed by Gallant and Tauchen (1996) (and used by Dai and Singleton (2000)), but the latter is substantially more computationally intensive.

For our specific application, the log-likelihood of the yield vector c is

$$L(c; \psi) = \frac{1}{T} \sum_{t=1}^T \left[\log p(v_t^\psi, r_t^\psi \mid v_{t-1}^\psi, r_{t-1}^\psi) - \log |\det Dg(v_t^\psi, r_t^\psi; \psi)| \right].$$

The maximum likelihood estimator, $\hat{\psi}$, for ψ is then defined by

$$\hat{\psi} = \arg \max_{\psi} L(c; \psi).$$

Under standard technical conditions (see, for instance, Davidson and MacKinnon (1993)), $\hat{\psi}$ is consistent and efficient with respect to the data (c_1, \dots, c_T) . In particular, as $T \rightarrow \infty$, $\hat{\psi} \xrightarrow{a.s.} \psi$ and

$$T^{1/2}(\hat{\psi} - \psi) \xrightarrow{\mathcal{L}} N(0, \Omega), \tag{9}$$

where $\xrightarrow{\mathcal{L}}$ denotes convergence in law, and

$$\Omega^{-1} = E \left(\frac{\partial L(c; \psi) \partial L(c; \psi)^\top}{\partial \psi \partial \psi} \right) = -E \left(\frac{\partial^2 L(c; \psi)}{\partial \psi \partial \psi^\top} \right). \tag{10}$$

Moreover, Ω can be estimated consistently, under these conditions, by the sample counterpart of (10).

The reference LIBOR swap term-structure model is estimated using weekly data on two- and ten-year swap rates for the period 1/1/1988 to 7/23/1999.

Based on a preliminary empirical investigation, we found that there was little deterioration in fit by imposing the constraints that $\lambda^v = 0$ (the market price of volatility risk is zero) and that, for a small constant $\varepsilon > 0$, $\theta^v = (1 + \varepsilon)/(2K^{vv})$. This ensures that the Feller condition, guaranteeing that v_t is strictly positive for all t with probability one, is satisfied with a margin of ε . Lastly, we set the long-run mean, θ^r , of the reference short rate, r , equal to its implied sample mean¹³ in order to improve the average fit of the reference curve to the swap rates across the relevant range of maturities used in our subsequent analysis of Russian bond spreads. The parameter estimates are given in Table IV in Appendix D.

For our purpose of constructing a reference curve relative to which Russian yield spreads can be investigated, the resulting fit seems more than adequate. Between maturities of two and ten years, the average pricing errors, measured as the average differences between historical and model-implied swap yields, are all less than five basis points, with sample standard deviations of the fitting errors of less than ten basis points. Even for swaps with fifteen years to maturity (used to compute model-implied credit spreads for the Russian bond with the longest maturity), the average fitting errors were under twenty basis points. This error is small relative to the magnitude of historical credit spreads in the Russian sovereign market.

C. A Parametric Model of Spreads

To accommodate idiosyncratic components in the short spreads of different bonds of the same issuer, we select one sovereign bond as a “benchmark,” and model the joint behavior of the riskless term structure and the benchmark short spread s as:

$$d \begin{bmatrix} v_t \\ r_t \\ s_t \end{bmatrix} = \begin{bmatrix} K^{vv} & 0 & 0 \\ K^{rv} & K^{rr} & 0 \\ K^{sv} & K^{sr} & K^{ss} \end{bmatrix} \left(\begin{bmatrix} \theta^v \\ \theta^r \\ \theta^s \end{bmatrix} - \begin{bmatrix} v_t \\ r_t \\ s_t \end{bmatrix} \right) dt + \begin{bmatrix} 1 & 0 & 0 \\ \Sigma^{rv} & \Sigma^{rr} & 0 \\ \Sigma^{sv} & \Sigma^{sr} & \Sigma^{ss} \end{bmatrix} \begin{bmatrix} \sqrt{v_t} \\ \sqrt{v_t} \\ 1 \end{bmatrix} dW_t, \quad (11)$$

where $W = (W^v, W^r, W^s)$ is a standard Brownian motion in \mathbb{R}^3 . The fact that s can take on negative values is not logically inconsistent with our theoretical model because of the possibility of a negative liquidity factor,

l , in (1).¹⁴ Extending (6), we take $\tilde{W} = (\tilde{W}^v, \tilde{W}^r, \tilde{W}^s)$ to be a risk-neutral standard Brownian motion with

$$dW_t^s = -\lambda^s dt + d\tilde{W}_t^s, \quad (12)$$

with λ^s being the market price of spread risk. Though not the most flexible parameterization of the market prices of risk (see Duffee (2000) and Dai and Singleton (2001)), this is relatively standard for affine models (Dai and Singleton (2000)).

For any non-benchmark bond $i \in \{1, \dots, I\}$ of the same issuer, we assume that its short-spread process is $s^i = s + \gamma^i$, where

$$d\gamma_t^i = \kappa^i (\theta^i - \gamma_t^i) dt + \sigma^i d\xi_t^i, \quad (13)$$

and ξ^i is a standard Brownian motion independent of $\{W, \xi^j : j \neq i\}$. The idiosyncratic factor γ^i is specific to bond i , and not to a maturity segment of the sovereign yield curve, and captures special liquidity or default features of the i^{th} bond. We experimented with alternative parameterizations of γ^i , including a square-root process, but this Gaussian formulation seemed to best fit the data.

The distribution of the idiosyncratic Brownian motion ξ^i under the risk-neutral measure is specified by

$$d\xi_t^i = (\kappa^i - \tilde{\kappa}^i) \gamma_t^i dt + d\tilde{\xi}_t^i, \quad (14)$$

where $\tilde{W}, \tilde{\xi}^1, \dots, \tilde{\xi}^I$ are independent standard Brownian motions under Q , and $\tilde{\kappa}^i$ is a constant. Following Duffee (2000), this specification offers the flexibility of different rates of mean reversion for γ under the actual and risk-neutral measures. This, in turn, implies that the risk premium associated with the idiosyncratic factor ξ_t^i contributes to time-variation in the excess return on the i^{th} bond, something that is precluded by the standard Gaussian specification, in which $(\kappa^i - \tilde{\kappa}^i) \gamma_t^i dt$ is replaced by $-\lambda^i dt$ in (14). Allowing this contribution of the Gaussian ξ to time-variation in excess returns improved the fit of our model.

To estimate the parameters

$$\phi = (\theta^s, K^{sv}, K^{sr}, K^{ss}, \Sigma^{sv}, \Sigma^{sr}, \Sigma^{ss}, \lambda^s, (\kappa^i, \theta^i, \sigma^i, \tilde{\kappa}^i)_{i=1, \dots, I})$$

governing the spreads, we use data on $I + 1$ (we have $I = 2$ in our application) defaultable bond prices, which we denote as b . The prices implied by the model are of the form $H(v_t, r_t, s_t, \gamma_t; \psi, \phi)$, where H is a sum

of known exponential-affine functions of the state vector $(v_t, r_t, s_t, \gamma_t)$, where $\gamma_t = (\gamma_t^1, \dots, \gamma_t^I)$. The state variables (r_t, v_t) are set to the values $(\hat{v}_t, \hat{r}_t) = (v_t^{\hat{\psi}}, r_t^{\hat{\psi}})$ implied by the reference model evaluated at the maximum-likelihood estimates $\hat{\psi}$. We then estimate ϕ by¹⁵

$$\hat{\phi} = \arg \max_{\phi} L(c, b; \hat{\psi}, \phi). \quad (15)$$

Under regularity, $\hat{\phi}$ is consistent and asymptotically normally distributed, with an asymptotic covariance matrix that (when applicable) reflects the “two-step” nature of the estimation strategy.

III. The Case of Russia

Using this model of sovereign credit spreads, we proceed to examine the pricing of Russian bonds during the late 1990’s. After a brief overview of the macroeconomic climate in Russia leading up to this default, we estimate our pricing model using weekly data over the sample period February 4, 1994 through August 7, 1998, ending just prior to the GKO default, which is the first clear credit event in the sample. We then explore the implications of fitting “through the default,” using the extended model (4).

A. Russian Debt and Macro-economic Developments

In this paper, we focus on the MinFins and the (dollar-denominated) Eurobonds, because there are natural reference curves relative to which we can study Russian credit spreads. Figure 2 shows the yield spreads of three MinFin bonds relative to US swap rates for our pre-default sample period from 2/4/1994 to 7/31/1998. The spreads of these bonds were high until the end of 1995, due to significant uncertainty regarding Russia’s ability and willingness to service its debt. In 1995, the Russian government implemented a series of reforms. IMF (1999b) writes that

“(. . .) the authorities resolved in 1995 to achieve a decisive reduction in inflation via a tight monetary policy supported by a halving of the federal budget deficit to under 6 percent of GDP. (. . .) Finally, to foster stability of the ruble, they adopted an exchange-rate band system from July 1995. Despite political pressures, the



Figure 2. Yield spreads (relative to the Dollar swap curve) of the Russian MinFin 3–5 bonds over the period 2/4/1994 to 7/31/1998.

authorities stuck to their program, and their main goals were achieved.”

Because of these reforms, the spreads of Russian MinFins started to decline in 1995. In July 1996, Yeltsin was re-elected, which increased trust in stability, and strengthened the market-oriented reformers, reducing spreads further. In this period, optimism was growing, and January-September 1997 is referred to as “The Zenith of Expectations” in IMF (1999b). In September 1997, the spread of the MinFin 3 reached an all-time low, just under 2 percent.

Nevertheless, the Russian government was running large budget deficits throughout the nineties, and the structural reforms implemented to improve the economic strength of the country were having limited success. (Russian real GDP contracted by an estimated 33% from 1993 to the end of 1998 (IMF (1999b)).) Starting in October 1997, and continuing into 1998, Russia’s problems started to further deteriorate (IMF (1999b)). The prices of Russian commodities — in particular, the price of Brent oil — declined. Partly as a

consequence, Russia's current account turned negative in 1997 for the first time since the start of the reforms. Furthermore, the Asian crisis was driving investors away from emerging markets globally, and in November 1997 the Ruble came under pressure, forcing Russia to increase domestic interest rates to defend the exchange-rate band. Hence, starting in October 1997, and continuing into the spring and summer of 1998, spreads were driven up, and Russia was downgraded by several credit-rating agencies.

These developments did not prevent Russia from issuing new Eurobonds — in fact Russia issued five Eurobonds in 1998, one in March, one in April, two in June, and one in July. The first four of these were oversubscribed (IMF (1999a)). Some market participants even talked about “feeding frenzies” in connection with new Russian issues. Yield-seeking investors evidently presumed that Russia was too important a country for major industrial nations to allow it to collapse. The 1998 Eurobonds were issued in anticipation of shortfalls in cash needed to make payments on short-term domestic debt and to service foreign debt. In addition, the IMF reluctantly agreed to give Russia additional loans. In August, 1998 it became clear that these measures were not enough to establish sufficient liquidity for Russia, and there was a substantial outflow of capital from Russia.

On August 17, 1998, Russia announced a compulsory restructuring of the domestic government debt (GKO and OFZ). A 90-day moratorium was placed on foreign commercial debt, and the exchange-rate band was devalued. At this time, it was unclear whether Russia would also default on external sovereign debt. Subsequently, however, Russia defaulted on the London-Club debt, the Paris-Club debt, and the principal of the MinFin 3, which matured in 1999. As of this writing, Russia remains current on interest payments on the remainder of the MinFins and Eurobonds.

Figure 1 shows the evolution of MinFin prices during and after this crisis. The prices of MinFins dropped initially to about 20–30% of their pre-default market values, but have since recovered substantially. The recovery can be attributed to Russia's favorable macroeconomic developments and to the fact that interest payments were made in both 1999 and 2000. Over the two years subsequent to this default, Russian real GDP grew at an annual rate of 6–7%, the current account was consistently positive due to reduced imports and increasing exports (largely driven by increasing oil prices), and the fiscal budget of the federal government was heading in 2000 toward a surplus (IMF (2000b)). Similarly, the Eurobond prices were reduced during the GKO default to around 30–40% of their pre-default market values, but

they too have largely recovered.

B. Pre-Default Analysis

Initially, we focus on pre-default weekly data, over the sample period 2/4/1994 to 7/31/1998, for the yields on the MinFin 3, 4, and 5 bonds. These are the dollar-denominated bonds with the longest available time series. MinFin 4 was chosen as the benchmark bond because it had the largest amount of principal outstanding (see Table I).

The *ML* estimates of the model are displayed in the first column of Table II. We have imposed four over-identifying restrictions because of the high

	2/94–7/98	SD	2/94–10/00	SD
θ^s	0.0200	(18.03)	0.1086	(31.45)
K^{sv}	0.0230	(0.017)	0.0125	(0.015)
K^{sr}	0.0898	(0.158)	0.1277	(0.160)
K^{ss}	0.2112	(0.019)	0.1367	(0.015)
κ^1	1.3433	(0.815)	1.2415	(0.767)
κ^2	1.8690	(0.961)	0.4699	(0.512)
Σ^{ss}	7.8895	(0.521)	10.867	(0.819)
σ^1	2.4126	(0.261)	2.4279	(0.254)
σ^2	0.3518	(0.189)	0.4261	(0.213)
λ^s	-0.2547	(0.483)	-0.2531	(0.398)
$\tilde{\kappa}^1$	-0.1989	(0.057)	-0.1825	(0.056)
$\tilde{\kappa}^2$	-0.2301	(0.064)	-0.2110	(0.062)

Table II: The first and third columns contain the parameter estimates for the sample periods 2/94–7/98 (the pre-default sample) and 2/94–10/00 (the full sample), respectively. The columns labeled *SD* give standard deviations estimated using the outer product of the first derivatives of the likelihood function. (Standard errors estimated using the second derivative of the likelihood function are similar.) These standard errors are not corrected for the facts that the swap parameters have been estimated, and that the Russian bond data are non-stationary because of changing maturities.

dimension of the parameter space. We set the long-run means, θ^1 and θ^2 , of the idiosyncratic factors, γ^1 and γ^2 , to zero, because they were found to be insignificantly different from zero at conventional levels of significance. Likewise, we set $\Sigma^{sv} = \Sigma^{sr} = 0$. The model allows the benchmark spread

factor, s , to be correlated with both the reference LIBOR short rate r and its volatility v , since K^{sv} and K^{sr} may be non-zero. We note that the idiosyncratic factors are estimated to be risk-neutrally “explosive” because $\tilde{\kappa}^1$ and $\tilde{\kappa}^2$ are negative. One interpretation of this is that investors have a high degree of risk aversion with respect to changes in these factors.

B.1. The Benchmark Spread

Figure 3 displays time series of the implied benchmark short spread (the instantaneous spread s for MinFin 4). The basic pattern in Figure 3 fol-

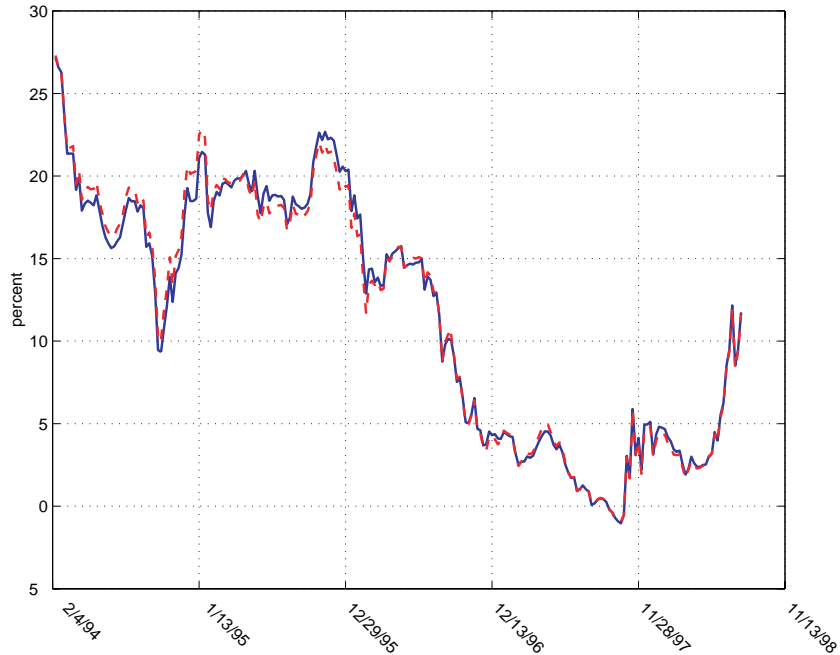


Figure 3. The solid line shows the benchmark short spread, s , and the dashed line shows the short spread adjusted for differences between the model-implied short rate and the 1-month LIBOR, that is $s_t + r_t - r_t^{obs}$.

lows that of the raw yields in Figure 2. It is notable, however, how low the short spread was in September and October, 1997. This may have been a manifestation of “over pricing” in the market, consistent with observations made by some market participants. For instance, Malleret, Orlova, and Romanov (1999) write: “In 1996 and 1997, Russia was the darling of inter-

national investors. ... the investors had been buying not the unappealing reality, but Russia’s expected bright future.” Additionally, the IMF (1999a) writes: “there is clear evidence that Russia represents a case where many investors bought securities that they did not fully understand ... Some investors in MinFins appear to have lacked a full understanding of their legal status, including the jurisdiction in which they were issued.” Without the benefit of such a dynamic term-structure model, it would have been difficult to pick up this puzzling pattern of the short-maturity MinFin and reference LIBOR instruments trading at comparable spreads.

Of course, we must also entertain the explanation of model misspecification or inaccurate parameter estimates. One potential problem could be that our model does not match LIBOR short rates perfectly. To determine the extent of this problem, we compute an adjusted short spread as $s_t + r_t - r_t^{obs}$, where r_t^{obs} is the 1-month LIBOR rate. Figure 3 shows that the difference between the short spread and its adjusted version is small. Indeed, the adjusted benchmark short spread is also negative when the short spread is. It may also be that the model is understating the short spread s during September, 1997. From Figure 6 we can infer that $s + \gamma^{MinFin3} > s$ during this period, so spreads on the shorter maturity MinFin 3 were larger than those on MinFin 4. However, it seems unlikely that such mis-specification would fully account for the strikingly low level of s , even if it reversed the finding that short-term MinFins would be trading through LIBOR (the period of negative s).

According to our macroeconomic overview, the risks of a Russian default or restructuring, and hence the short spread s , should be influenced by government balance-sheet variables and related variables that influence internal and external balances. One such variable is the price of Brent oil. Fuel products constituted between 40% and 50% of Russia’s exports in the period that we consider, and oil products contributed to more than half of this (IMF (1999b)). Hence, the oil price is important in determining both Russia’s current account and budget deficit. Another relevant credit-related variable is the total level of foreign currency reserves, minus gold, held by the Russian central bank.

As a preliminary examination of the potential influence of these variables, we regressed the fitted spread s onto the Brent oil price index (OIL) and total Russian reserves ($RESERVE$) using monthly data, and obtained

$$s_t = 26.7 - 0.277 * OIL_t - 0.0010 * RESERVE_t + \varepsilon_t,$$

(8.9) (0.52) (0.0002)

with an R^2 of 81.4%. Standard errors, corrected for heteroscedasticity using the Newey-West method, are shown parenthetically. As expected, both regression coefficients are negative — the spread is low in periods when the oil price is high and the reserves are high — although only the *RESERVE*-coefficient is statistically significant at conventional levels.

To explore the dynamic relations among these variables, we estimate a sixth-order vector-autoregression (VAR) for *RESERVE*, *OIL*, and s , in that order. Figure 4 displays the implied orthogonalized impulse-response functions, along with two-standard-error confidence bands computed by Monte Carlo using the approach outlined in the RATS regression package manual. We see that shocks to the spread itself die out during a few months. Shocks to the oil price and reserves, on the other hand, build up over almost a year. As we would expect, positive shocks to oil prices and reserves are estimated to imply a reduction in the spread in the following period.

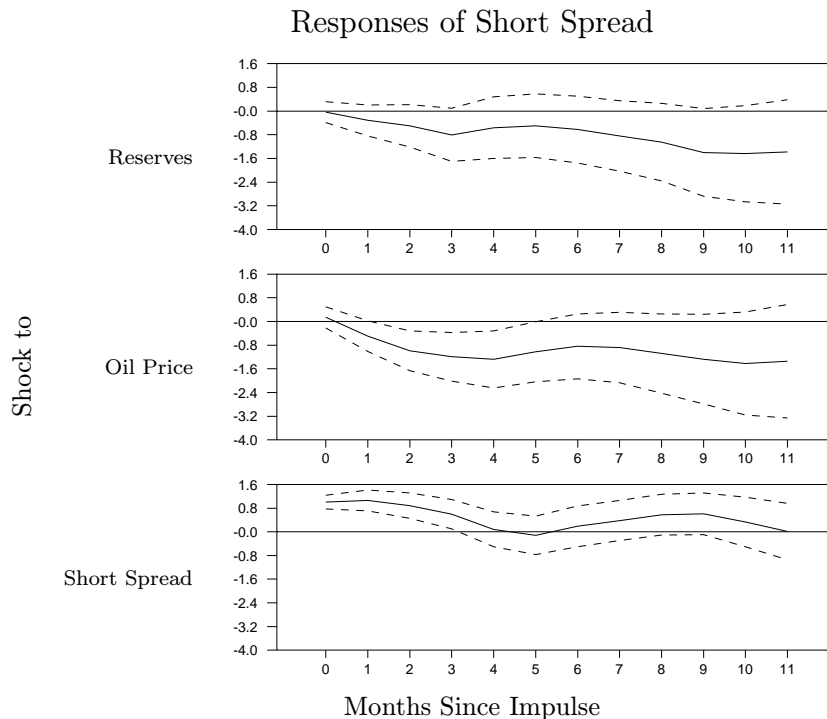


Figure 4. Orthogonalized impulse-response functions of the short rate based on a VAR of the short rate, the oil price, and Russian reserves.

Using the estimates of the benchmark short-spread process, we can also decompose the term structure of model-implied credit spreads on Russian zero-coupon bonds into components associated with the risk factors r , v , and s . The model implies that the yield of a reference zero-coupon bond with maturity τ is of the form $A(\tau) + B^v(\tau)v_t + B^r(\tau)r_t$, and that the yield of a zero-coupon bond with the same credit risk as the MinFin 4 is of the form $\bar{A}(\tau) + \bar{B}^v(\tau)v_t + \bar{B}^r(\tau)r_t + \bar{B}^s(\tau)s_t$, where A , B^v , B^r , \bar{A} , \bar{B}^v , \bar{B}^r , \bar{B}^s can be found as solutions to Riccati equations (Duffie and Kan (1996)). Hence, the term structure of zero-coupon yield spreads at time t is given, at maturity τ , by,

$$\bar{A}(\tau) - A(\tau) + (\bar{B}^v(\tau) - B^v(\tau))v_t + (\bar{B}^r(\tau) - B^r(\tau))r_t + \bar{B}^s(\tau)s_t. \quad (16)$$

The estimated mean term structure of yield spreads, found by replacing the state variables in (16) by their model-implied sample means, is downward sloping as is common for low-quality bonds.

Each “triple” of lines in Figure 5 shows each term of (16), with the coefficients replaced by their estimates and the state variables replaced by their implied sample means (middle figure), as well as means plus and minus twice the sample standard deviations of the state variables (corresponding to the upper and lower dashed lines). We see that changes in the level of the instantaneous reference short rate have almost no effect on the term structure of yield spreads, for given values of v and s . When v increases, the spreads on long-maturity bonds narrow, though variation in v contributes only marginally to the overall variation in the spread. Most of the variation of the zero spreads is due to changes in s (with possible indirect effects of r and v through their correlations with s). The sample correlation between the changes in r and in s is -0.18 , which is consistent with the negative correlations between instantaneous reference and spread factors in empirical studies of corporate bond yields (for instance, Duffee (1999)). The sample correlation of changes in v and s is 0.46 . Since changes in v have a correlation with changes in the ten-year swap yield of approximately 0.90 , the positive correlation of v and s suggests that movements in the long end of the swap curve are positively correlated with changes in short-term sovereign credit spreads.

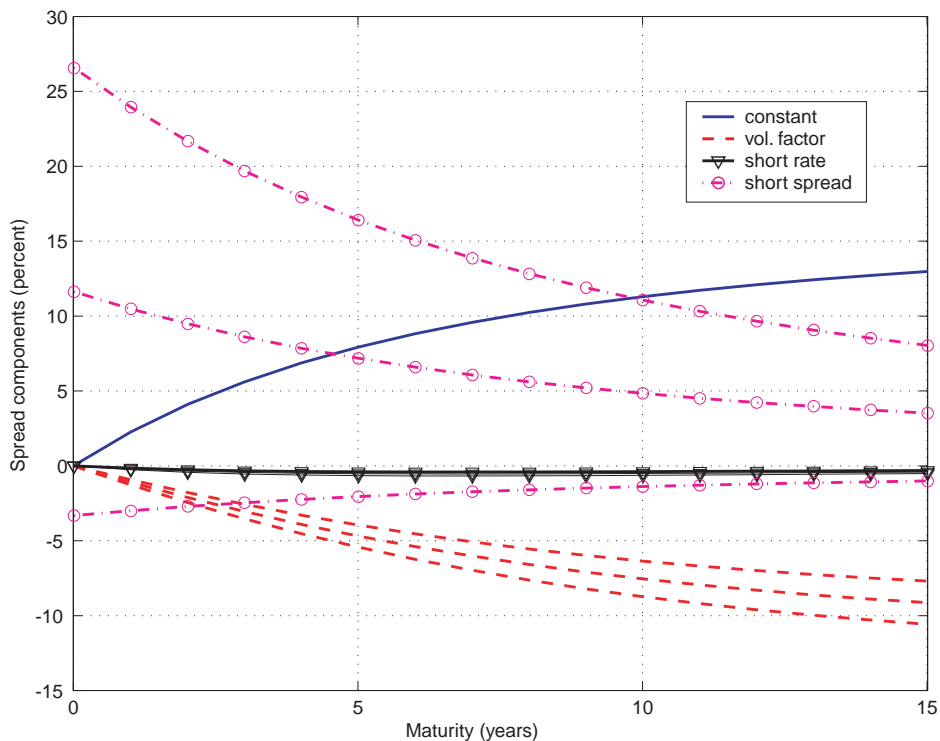


Figure 5. A decomposition of the term structure of spreads into its components. Each “triple” of lines shows each term of (16) evaluated at different values of the state. The (upper, middle, lower) line is obtained with each state variable replaced by $(\mu + 2\sigma, \mu, \mu - 2\sigma)$, where μ represents its mean and σ represents its volatility.

B.2. Relative Spread Behavior

Including the idiosyncratic factor γ^i in the discount rates of a non-benchmark MinFins implies that the fitted MinFin yields match their sample counterparts exactly. The model-implied estimates of these factors are plotted in Figures 6 and 7. If the γ^i processes are near zero, the markets for the three MinFins are well integrated and they share expected recoveries and default likelihoods. On the other hand, large absolute values of γ^i would indicate significant differences in credit characteristics or liquidity between the i^{th} bond and the benchmark. Clearly, during several periods, the market was treating MinFin 3, in particular, differently than MinFin 4.

Data limitations make it difficult to determine a breakdown of the total

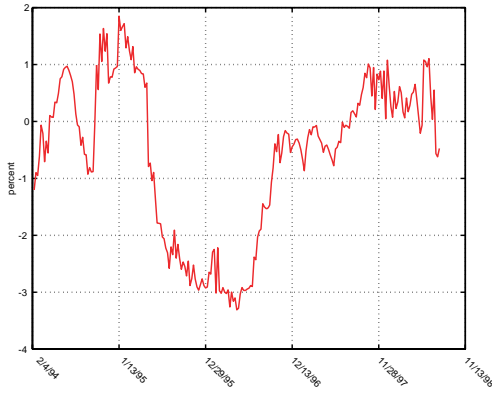


Figure 6. Idiosyncratic factor, γ^1 , for MinFin 3.

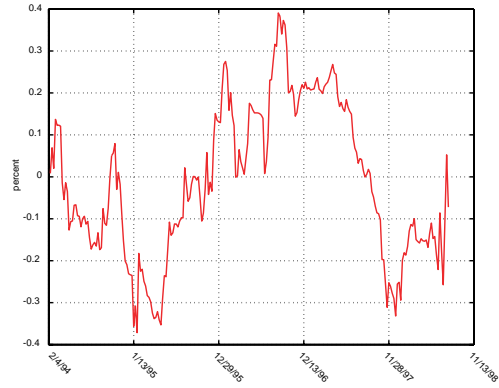


Figure 7. Idiosyncratic factor, γ^2 , for MinFin 5.

idiosyncratic factor γ^i into perceived differences (from MinFin 4) in liquidity, expected recovery, or the likelihood of a default or restructuring. In an attempt to assess the effects of relative liquidities on each γ^i , we compared the time paths of the “relative bid-ask spread” (the bid-ask spread divided by the mid price) to those of the fitted γ^i . There did not appear to be strong co-movements induced by changes in the bid-ask spreads, however, so we are led to conclude that relative liquidities as measured by bid-ask spreads are not the primary explanation for the variation in idiosyncratic spreads of MinFins 3 and 5.

Indirect evidence of the contributions of credit risk to γ^i comes from aligning movements in these idiosyncratic spread factors with key macroeconomic developments. We focus on the idiosyncratic factor, γ^1 , of MinFin 3, since it is larger in magnitude than that of MinFin 5. Initially, γ^1 varies around zero. Then, from May 1995 to August/October, 1996, γ^1 is negative. After October 1996, it is again close to zero. One story for this pattern is that the Russian reforms implemented in 1995 were viewed positively by the market, but were expected to have mostly shorter-term effects. Therefore, the shortest bond, MinFin 3, became expensive (compared to MinFin 4), leading to the period of negative γ^1 . Then, in late 1996, after Yeltsin’s election, the market accepts these improvements as more permanent, and the MinFin 3 aligns with the MinFin 4. An altogether different, non-credit-based, interpretation is that investor euphoria about Russia first showed up in the prices of MinFin 3’s in early 1996, perhaps because of ownership distribution or

clientele effects. It is also possible that the idiosyncratic factors are driven by model misspecification, or perhaps differences in the degree to which these bonds are on special in repo markets. The fact that γ^1 and γ^2 are somewhat negatively correlated suggests that it might be desirable to have an additional factor for the (benchmark) spread model to describe the slope of the term structure of spreads.

Another way of looking at differences in spreads is to consider the *ratios* of the short spreads associated with different bonds. For instance, the ratio between the MinFin 3 short spread and the benchmark (MinFin 4) short spread is:

$$\frac{s + \gamma^1}{s} = \frac{h^{\text{MinFin 3}} L^{\text{MinFin 3}} + l^{\text{MinFin 3}}}{h^{\text{MinFin 4}} L^{\text{MinFin 4}} + l^{\text{MinFin 4}}}. \quad (17)$$

If illiquidity is not priced or is a relatively unimportant factor in the volatility of $s + \gamma^1$ (that is, $l^{\text{MinFin 3}} \approx l^{\text{MinFin 4}} \approx 0$), then the ratios of instantaneous spreads can be interpreted as ratios of risk-neutral *mean loss rates* (within our recovery-of-market-value framework). If, in addition, these two bonds experience credit events at the same times, then their risk-neutral default arrival intensities are identical and, hence, the ratio (17) is the ratio of the respective risk-neutral expected fractional losses at default, $L^{\text{MinFin 3}}/L^{\text{MinFin 4}}$.

Figure 8 shows the fitted ratios (17) for the MinFin 3 and 5, omitting the period during which the levels of the $s + \gamma^i$ were close to zero. Under the last interpretation, we see that the relatively “rich” pricing of MinFin 3 compared to MinFin 4 during 1996 and the first half of 1997 could be attributed to lower expected loss at default on the MinFin 3 (about 80%–90% of the MinFin 4’s expected default loss). Along this line of interpretation, the expected loss the MinFin 5 at default is approximately 110% of the MinFin 4’s expected loss in mid 1997, but the ratio drops to around 90% by the end of 1997. Alternatively, if one assumes that the bonds have the same expected fractional loss at default, then these spread ratios can be interpreted as the ratios of the risk-neutral default intensities ($h^{\text{MinFin 3}}/h^{\text{MinFin 4}}$).

C. Analysis “Through the Default”

In this section, we investigate the properties of bond yields over our longer sample period that includes the 1998 default. Using our model, we compute investor’s expectations about the severity of the 1998 default. Given the limitations of the available Russian data and the single default experience,

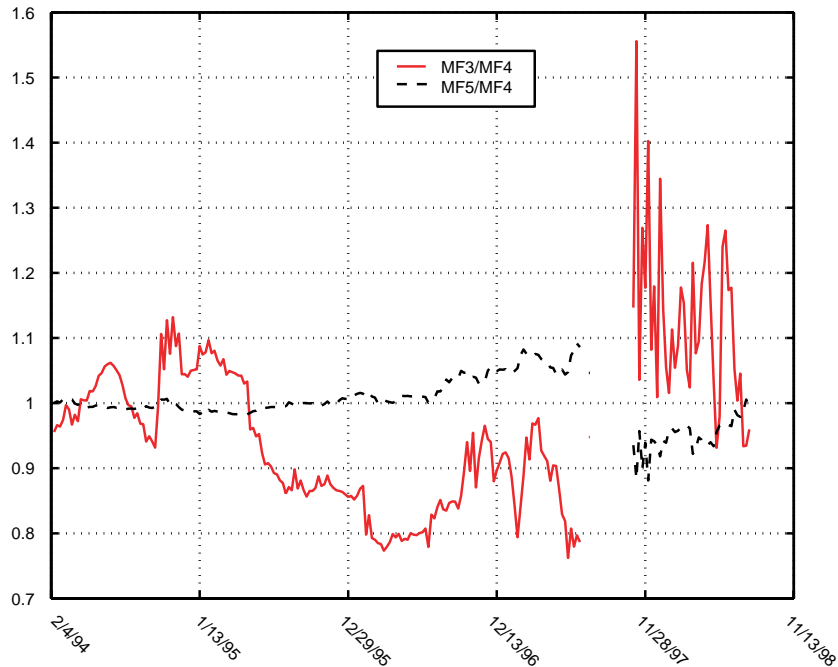


Figure 8. Ratios of short spreads.

we are led to make several strong assumptions that help in solving important identification and estimation problems. Therefore, we proceed in the spirit of a “calibration” rather than with formal *ML* estimation.

In order to capture the features of data discussed in Section I.A, we shift our focus to the extended pricing relation (4), allowing for uncertainty after the credit event for the eventual write-downs. In this formulation, the risk-neutral expected write-down, $E_t^Q(Y_n)$, associated with the n^{th} credit event may be bond-specific, so we would in principle need to consider an additional state variables for each bond, at each of the N_T credit events. In the case of Russia, $N_T = 1$, but this still leaves us with two state variables for each bond yield (γ_i and $E_t^Q(Y_n^i)$), making it infeasible to invert the model for the implied states.

To overcome this identification problem for our analysis of Russian MinFins, we make the simplifying assumptions that (i) the MinFin 4 and MinFin 5 bonds have the same short spread in the post-default period; (ii) the risk-neutral expected write-downs of these bonds are perfectly correlated; and

(iii) as coupon payments are made, they reveal no information relevant to risk-neutral conditional expected recoveries on subsequent coupons and principals. We focus on the MinFin 4 and MinFin 5 bonds because the MinFin 3 matured in May, 1999. The first assumption is largely consistent with our pre-default analysis, where we found that these bonds had similar short spreads ($\gamma^{\text{MinFin 5}}$ was small). The second assumption allows us to express the expected write-down of MinFin 5 as an affine function, say η , of the expected write-down of the MinFin 4, that is, $E_t^Q[Y_1^{\text{MinFin 5}}] = \eta(E_t^Q[Y_1^{\text{MinFin 4}}])$. The affine function η is then determined by the following two conditions: (i) $\eta(1) = 1$, that is, if one of these bonds fully recovers, then so does the other, and (ii)

$$\eta\left(\frac{b^{\text{MinFin 4}}(8/21/1998)}{b^{\text{MinFin 4}}(8/7/1998)}\right) = \frac{b^{\text{MinFin 5}}(8/21/1998)}{b^{\text{MinFin 5}}(8/7/1998)},$$

where $b^i(t)$ denotes the price of bond i at time t . Condition (ii) is motivated by supposing that the price drops observed in August 1998 reflected investor's expected write downs. Then we "observe" the expected write downs of MinFins 4 and 5 at that time, and these "observations" imply the restriction on η that is given by (ii).

With these assumptions, the prices of MinFin 4 and 5 are known functions of their common short spread, s_t , and the expected write down, $E_t^Q[Y_1^{\text{MinFin 4}}]$, of MinFin 4. Hence, for each date we can compute the model-implied short spread and risk-neutral expected write downs corresponding to the observed prices by solving two equations with two unknowns. For simplicity, we henceforth denote $Y_1^{\text{MinFin 4}}$ by Y .

Another consideration is the nature of the conditional risk-neutral expected write-down process U , defined by $U_t = E_t^Q(Y)$. The fact that U is a risk-neutral martingale bounded between zero and one suggests a parameterization under which U is not stationary. For example, we could suppose that

$$U_t = e^{-\kappa^z(T-t)}(Z_t - \theta^z) + \theta^z, \quad (18)$$

where Z is the Jacobi diffusion, defined by

$$dZ_t = \kappa^z(\theta^z - Z_t) dt + \sigma^z \sqrt{Z_t(1 - Z_t)} d\tilde{W}_t^z, \quad (19)$$

and where \tilde{W}^z is a standard Brownian motion under Q . So long as $2\kappa^z\theta^z > (\sigma^z)^2$ and $2\kappa^z(1 - \theta^z) > (\sigma^z)^2$, Z_t has support $(0, 1)$. Although Z has a

risk-neutral stationary distribution,¹⁶ U_t is non-stationary through its dependence on $(T - t)$. Moreover, to construct the likelihood of the data, we need the actual conditional density of Z_t . This density might reasonably be assumed to depend on (r_t, v_t, s_t) under the actual probability measure, in contrast to the parameterization (18)-(19). (Recall that, for pricing, we assume independent recovery only under the risk-neutral measure Q .)

Rather than attempting to parametrically model the (conditional) dependencies among $(r_t, v_t, s_t, N_t, Z_t)$ with our limited data, we instead proceed to “calibrate” ϕ to the data by maximizing the “loss” function

$$CL(c, b; \hat{\psi}, \phi) = \left[\prod_{t=1}^{\tau-1} p(s_t | X_{t-1}^{(1)}, v_t, r_t; \phi) p(\gamma_t | \gamma_{t-1}; \phi) \frac{1}{|\det DH_s(X_t^{(1)})|} \right] \cdot \left[\prod_{t=\tau}^T p(s_t | X_{t-1}^{(1)}, v_t, r_t; \phi) \frac{1}{|\det DH_{s,Z}(X_t)|} \right]. \quad (20)$$

The loss function (20) is obtained from the actual likelihood function by leaving out all terms involving $p(N)$ and $p(Y)$.¹⁷ This choice is motivated by its close resemblance to the partial likelihood function for ϕ obtained under the additional assumption of *Independent Recovery* under the actual probability measure.¹⁸ The latter would be an even stronger assumption than we have made for pricing, so we view CL as a calibration function and not a likelihood function.

We use weekly data on MinFins 3–5 from February 4, 1994 until August 7, 1998, and weekly data on MinFins 4–5 from August 21, 1998 through September 8, 2000. The estimated parameters in column three of Table II are relatively similar to those estimated using pre-default data only. A notable difference is that the estimate of the volatility, Σ^{ss} , of the short spread is much larger in the full sample.

C.1. Model-Implied Short Spread and Write-Down

Figure 9 shows the implied short spread s over the full sample period. This spread seems to be more volatile after the 1998 crisis than before. If true, this would be inconsistent with the stationarity assumption underlying our econometric analysis. The first large upward spike in s occurs shortly after 1998 credit event, while the second spike occurs around the time of the 1999 dividend payments and the default on MinFin 3 (May, 1999). Interestingly,

the market prices of the MinFins 4 and 5 do *not* change substantially around the time of the second spike in s , an observation to which we will return.

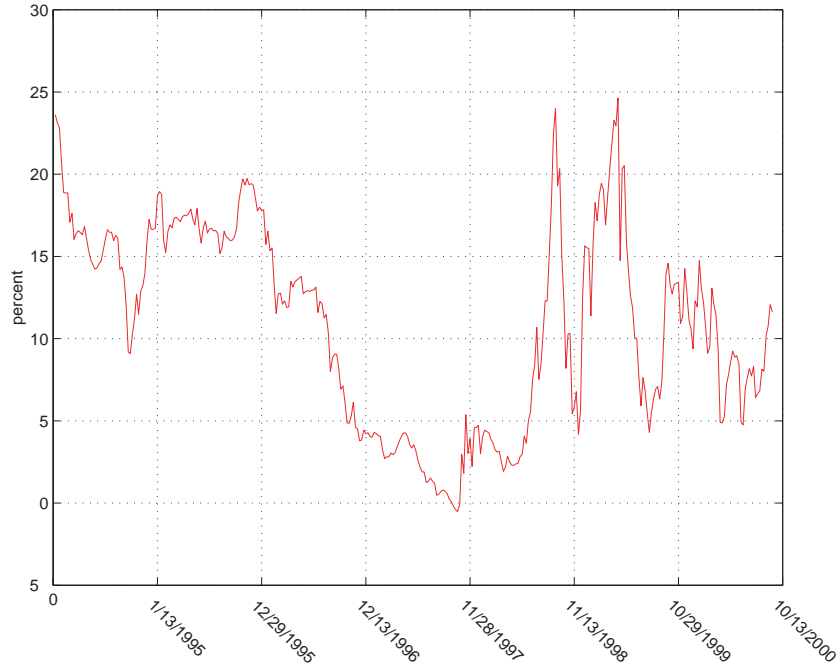


Figure 9. Short spread on the “expected payment” of the MinFin 4.

Figure 10 shows the model-implied expected write-downs of the MinFin 4 and MinFin 5 bonds. We see that the expected write-downs increase dramatically starting in early 1999. Hence, the model implies that investors became more optimistic about the consequences of the 1998 event through 1999 and 2000. The rising optimism coincided with such favorable economic developments as increased Russian profitability, an improved trade balance, and rising energy prices (see Section A), and by a seemingly higher than expected willingness of the Russian government to make interest payments. For example, in May 1999 when Russia decided to pay interest on the MinFins, a Paribas analyst stated “It was a positive surprise — no one expected to get any money in May,” and an analyst at Lehman Brothers said: “This is clearly good news. And it runs contrary to the belief in the market that the threes, fours, and fives would be restructured” (*The Russia Journal*, Issue 15).

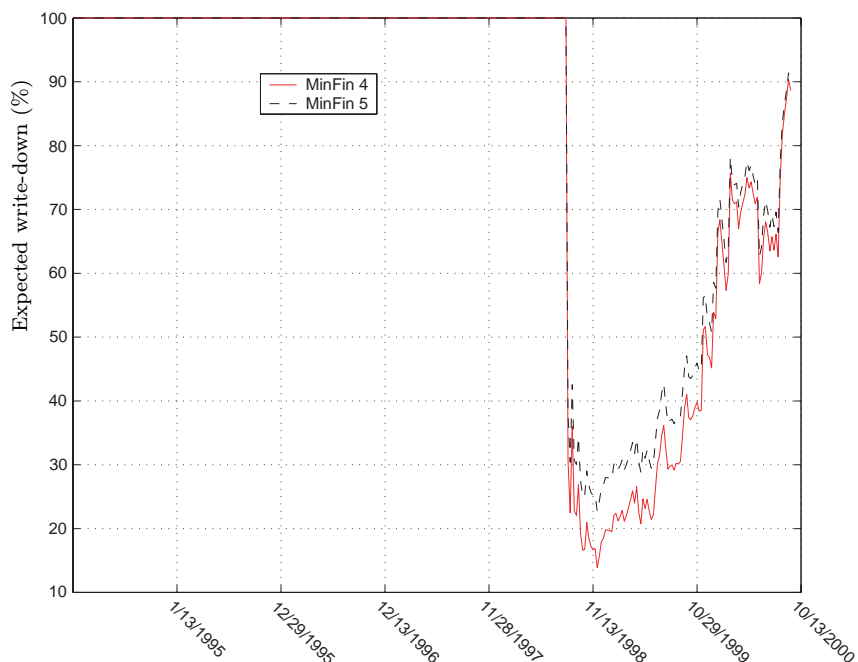


Figure 10. Model-implied risk-neutral conditional expected write downs, MinFin 4 and MinFin 5 bonds.

D. Pricing Out-of-Sample Bonds

Figures 11 and 12 show the market and model-implied prices of two Russian bonds that were not used in estimation, MinFin 6 and Euro-01 (see Table I for contract terms).¹⁹ We see that the model prices that are based on the pre-default and full-sample parameters, respectively, are similar, which is perhaps not surprising since the parameters are similar. Hence, we focus on the differences between model prices and observed prices.

The models fit the MinFin 6 well until the credit event of August, 1998. This means that the market did not differentiate between the MinFins 4 and 6 during this time period. Immediately after the crisis, the model marginally underprices the MinFin 6. This underpricing increases dramatically in the week of 4/23/1999. On 4/20/1999, First Deputy Finance Minister Mikhail Kasyanov made clear that Soviet-era debt would take much lower priority than post-Soviet debt and that, with regard to the MinFins, Issues 3 through 5 were considered Soviet-era, while Issues 6 and 7 were considered post-Soviet

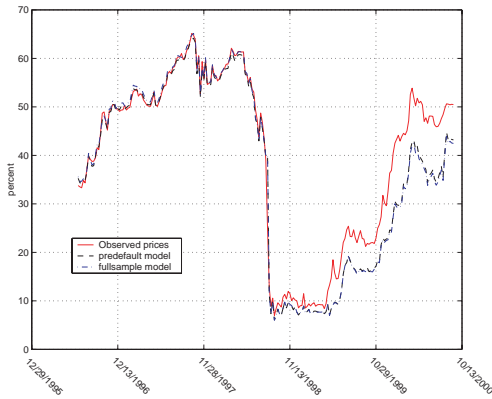


Figure 11. Empirical and model-implied prices of MinFin 6.

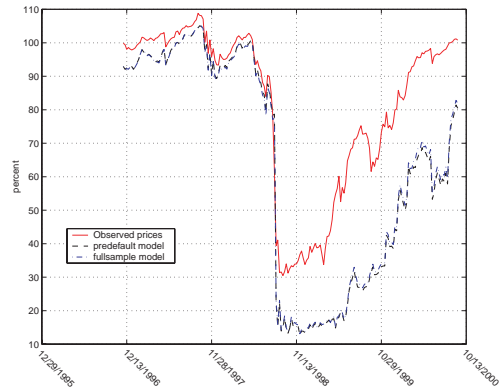


Figure 12. Empirical and model-implied prices of Euro-01.

debts (*The Russia Journal*, Issue 13). Kasyanov further stated that Russia would not pay the principal on MinFin 3 at its due date 5/14/99, but rather would postpone the payment. Consistent with these observations, MinFin 6 becomes “rich” relative to prices based on the MinFin 4 model.

Recall that the prices of MinFins 4 and 5 did not change substantially around these announcements. The relative stability of their prices seems to reflect two countervailing forces: (i) heightened fears of future credit events for “Soviet-era” MinFins raised the short spread s , and (ii) improving economic conditions in Russia increased expected redemption values relative to the more pessimistic expectations at the time of the August, 1998 default.

Euro-01 consistently traded at prices around or above the model-implied prices, suggesting that market participants have always viewed the Russian Eurodollar bonds as at least as good or better than the MinFins. The model-implied prices of this Eurobond are lower than the observed prices prior to September 1997, match relatively well the observed prices from September 1997 until August, 1998, and are substantially lower than Eurobond prices after August, 1998. This is consistent with changing market perceptions of the relative likelihoods that Russia would default on (i) its global debt only, or (ii) its internal debt only. MinFin bonds are considered domestic bonds and are partly held by residents of Russia, whereas Eurobonds are global bonds held abroad. Russia may be viewed as more likely to default on internal debt when the country as a whole is productive and needs a good relationship with other countries in order to attract capital and encourage trade. Russia

may be viewed as more likely to default on external debt when the country is doing so poorly that it is unlikely to attract much capital anyway, and the government may decide that making payments to residents is more important than to foreign institutions. Also, political favors can be particularly important in such poor times. After Yeltsin's victory in 1996 and until September 1997, there was a high degree of confidence in Russia's political stability and economy. This is perhaps why investors expected case *(ii)* to be the more likely scenario, therefore causing Eurobonds to trade at lower spreads than MinFins. When the Russian economy's problems mounted in late 1997 and 1998, investors re-evaluated the Russian risk — they became unsure about whether case *(i)* or case *(ii)* was more likely. This story is consistent with the fact that MinFins and Eurobonds were trading at similar spreads during this period. In August, 1998, Russia defaulted on its Ruble-denominated Treasuries. This, according to this story, convinced investors that case *(ii)* was the more likely one, and MinFin prices dropped more severely than did Eurobond prices.

Appendix A: Proof of Theorem 1

The proof of this theorem uses standard techniques and is only sketched. The discounted gain process is

$$G_t = - \int_0^t e^{-\int_0^v r_u du} l_v P_v dv + \int_0^t e^{-\int_0^v r_u du} W_{N_v} P_{v-} dN_v + e^{-\int_0^t r_u du} P_t.$$

The first term is the discounted costs of illiquidity. The second term is the discounted cash flows generated at the times of credit events, which are the times at which N jumps by 1. The third term is the discounted price. With the proposed price process P stated in the theorem, the final price is $P_T = X \prod_{n \leq N(T)} Y_n$, as required. After an application of Itô's Formula, as in Duffie and Singleton (1999), the discounted gain process is a Q -martingale, using the fact that $N_t - \int_0^t h_u du$ is a Q -martingale. These properties of P and G are necessary, from the definition of Q as an equivalent martingale measure, for P to be the price process. There is a unique price process with this property, as explained in Duffie and Singleton (1999).

Appendix B: Transition Density

In this appendix, we derive an approximation of the transition density of an affine diffusion of the form²⁰ $X = (X^I, X^D) \in \mathbb{R}_+^n \times \mathbb{R}^m$, where X^I is an affine vector process driving the volatility, in the sense that X solves the stochastic differential equation:

$$\begin{aligned} dX_t^I &= (k^I - K^{II} X_t^I) dt + \sqrt{\Delta(X_t^I)} dW_t^I, \\ dX_t^D &= (k^D - K^{DI} X_t^I - K^{DD} X_t^D) dt + \sqrt{\Delta(\alpha^D + B^{DI} X_t^I)} dW_t^D, \end{aligned}$$

where $\Delta(x)$ denotes a diagonal matrix whose diagonal elements are the respective elements of the vector x , and where (W^I, W^D) is a standard Brownian motion in $\mathbb{R}^n \times \mathbb{R}^m$. We wish to evaluate the transition density $p(X_\delta \mid X_0)$, where δ is the time between consecutive observations. We take advantage of Bayes' Rule, and the fact that X^I is itself a Markov process, to obtain

$$\begin{aligned} p(X_\delta \mid Y_0) &= p(X_\delta^I \mid X_0) p(X_\delta^D \mid X_0, X_\delta^I) \\ &= p(X_\delta^I \mid X_0^I) p(X_\delta^D \mid X_0, X_\delta^I). \end{aligned} \tag{21}$$

Our method is based on the assumption that the matrix K^{II} is diagonal, or equivalently that the elements of X^I are independent. In our two-factor model of the riskless term structure, this assumption is trivially satisfied, since X^I is the one-dimensional process v . If $n > 1$, however, this independence assumption represents a restriction on the state process that, if violated, means that the transition density of X^I is not known in closed form. Under this independence assumption, the conditional distribution of X_δ^I given X_0^I is a product of non-central chi-squares (Feller (1951)), with a density given by

$$p(X_\delta^I \mid X_0^I) = \prod_{i \in I} p(X_\delta^i \mid X_0^i), \quad (22)$$

where

$$p(X_\delta^i \mid X_0^i) = c_i e^{c_i(X_\delta^i + e^{-K_{ii}\delta} X_0^i)} \left(\frac{X_\delta^i}{e^{-K_{ii}\delta} X_0^i} \right)^{q_i/2} \quad (23)$$

$$\cdot I_{q_i} \left(2c_i (X_\delta^i X_0^i e^{-K_{ii}\delta})^{1/2} \right), \quad (24)$$

with $c_i = 2K_{ii} (1 - e^{-K_{ii}\delta})^{-1}$, $q_i = 2k_i - 1$, and I_q denotes the modified Bessel function of the first kind of order q .

There is no known explicit expression for $p(X_\delta^D \mid X_0, X_\delta^I)$. We base an approximation on the following observation: The distribution of X_δ^D conditional on X_0^D and the entire path of X^I from time 0 to time δ has a known normal density

$$p(X_\delta^D \mid X_0, \{X_s^I, s \in [0, \delta]\}) = \varphi(X_\delta^D, m_\delta, V_\delta), \quad (25)$$

where $\varphi(\cdot, m, V)$ is the density of a normal with mean m and variance V ,

$$\varphi(x, m, V) = ((2\pi)^N |\det(V)|)^{-\frac{1}{2}} e^{-\frac{1}{2}(x-m)^\top V^{-1}(x-m)}, \quad (26)$$

and

$$m_\delta = e^{-K^{DD}\delta} \left[\int_0^\delta e^{K^{DD}s} (k^D - K^{DI} X_s^I) ds + X_0^D \right] \quad (27)$$

$$V_\delta = e^{-K^{DD}\delta} \left[\int_0^\delta e^{K^{DD}s} \Delta(\alpha^D + B^{DI} X_s^I) e^{K^{DD^\top}s} ds \right] e^{-K^{DD^\top}\delta}. \quad (28)$$

Hence, by the law of iterated expectations,

$$p(y_\delta^D \mid X_0, X_\delta^I) = E(p(y_\delta^D \mid X_0, X_s^I, s \in [0, \delta]) \mid Y_0, Y_\delta^I) \quad (29)$$

$$= E(\varphi(y_\delta^D, m_\delta, V_\delta) \mid X_0, X_\delta^I). \quad (30)$$

Thus, completing the specification of the conditional density function of the state amounts to approximating the expectation in (30). We can approximate $p(\cdot \mid X_0, X_\delta^I)$ as the conditional density of X_δ^D given X^I , evaluated at an outcome of the path of X^I that is linear between X_0^I and X_δ^I . We have found this approximation to be tractable and accurate for our application, as is shown below.

We can approximate the expectation in (30) with arbitrary accuracy, however, at the expense of some tractability. This can be done by simulating the path of X^I conditional on its starting point, X_0^I , and its endpoint, X_δ^I . In particular, we divide the interval $[0, \delta]$ into J sub-intervals, and let X^J be the piecewise-linear approximation to X^I corresponding to this division of the interval. (We could use other interpolation schemes.) We define m^J and V^J similarly to m and V , respectively, but with X^I replaced with X^J . This is like approximating the integrals in (27) and (28) with Riemann sums. Hence, we have (a.s.) that $(m_\delta^J, V_\delta^J) \rightarrow (m_\delta, V_\delta)$, as $J \rightarrow \infty$. Using this result as well as dominated convergence, under technical regularity²¹ we have

$$p(y_\delta^D \mid X_0, X_\delta^I) = E(\varphi(y_\delta^D, m_\delta, V_\delta) \mid X_0, X_\delta^I) \quad (31)$$

$$= E\left(\lim_{J \rightarrow \infty} \varphi(y_\delta^D, m_\delta^J, V_\delta^J) \mid X_0, X_\delta^I\right) \quad (32)$$

$$= \lim_{J \rightarrow \infty} E(\varphi(y_\delta^D, m_\delta^J, V_\delta^J) \mid X_0, X_\delta^I) \quad (33)$$

$$= \lim_{J \rightarrow \infty} p^J(y_\delta^D \mid X_0, X_\delta^I), \quad (34)$$

where p^J is defined by the last equation. This gives us a sequence, $\{p^J\}$, of approximations to the density, that converges to the true density. The J 'th approximation, p^J , can be computed by Monte Carlo integration:

$$p^J(y_\delta^D \mid X_0, X_\delta^I) = \lim_{Z \rightarrow \infty} \frac{1}{Z} \sum_{z=1}^Z \varphi(y_\delta^D, m_\delta^{J,z}, V_\delta^{J,z}), \quad (35)$$

where $(X^{J,z})_{z=1,2,\dots}$ are independent simulations of $(X_0^I, X_{\delta/J}^I, \dots, X_\delta^I)$, drawn from their joint conditional distribution given X_0^I and X_δ^I . We take $m_\delta^{J,z}$ and

$V^{J,z}$ to be the corresponding mean and variance, based on linear interpolations. We can compute $m^{J,z}$ and $V^{J,z}$ explicitly since we need only compute integrals of exponentials. We know the transition densities of X^I in closed form, but simulating from this density is not trivial. Instead, we can simulate from another (possibly incorrect) density for $X^{J,z}$, say $\hat{f}(\cdot)$, and use “importance sampling” to compute p^J as

$$p^J(y_\delta^D \mid X_0, X_\delta^I) = \lim_{Z \rightarrow \infty} \sum_{z=1}^Z \varphi(y_\delta^D, m_\delta^{J,z}, V_\delta^{J,z}) w^{z,Z}, \quad (36)$$

where $(X^{J,z})_{z=1,2,\dots}$ are independent draws from $\hat{f}(\cdot)$, $f(\cdot)$ denotes the true density of $(X_0^I, X_{\delta/J}^I, \dots, X_\delta^I)$ conditional on X_0^I and X_δ^I , and

$$w^{z,Z} = \frac{f(X^{J,z})/\hat{f}(X^{J,z})}{\sum_{l=1}^Z f(X^{J,l})/\hat{f}(X^{J,l})}. \quad (37)$$

The first-order ($J = 1$) density approximation p^1 *does not require simulation*, since it is based on the (linear) interpolation between X_0^I and X_δ^I . Next, we consider the accuracy of this approximation for our application.

Accuracy of Approximation

To assess the accuracy of our approximation, we compute $p^i(y_\delta^2 \mid X_0^1, X_0^2, X_\delta^1)$ over a range of values of y_δ^2 , for $i = 1, 2, 3, 6$. We use the parameter estimates from our 2-factor swap model with the state space rotated to make Σ diagonal. We fix X_0^1 and X_0^2 at their long-run means, X_δ^1 at its long-run mean plus its weekly standard deviation, and let δ correspond to a weekly time step. The density approximations are computed using (36) with 40,000 simulations. (Computing p^1 does not require simulation.)

Figure 13 shows that the resulting first- and sixth- order approximate conditional densities are virtually on top of each other. Also displayed is 10^4 times the difference between p^1 and p^2 . This difference is partly due to approximation errors (related to p^1), and partly due to simulation noise related to p^6 . The corresponding results for the logarithm of the conditional densities are displayed in Figure 14.

Table III presents 10^4 times the average differences between p^i and p^6 for $i = 1, 2, 3$, over the range plotted in Figures 13 and 14. These differences are small both in absolute and relative terms, and both for densities and log-densities.

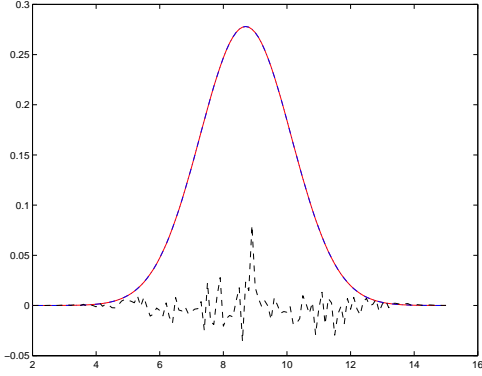


Figure 13. The first-order approximation, $p^1(\cdot \mid X_0, X_\delta^I)$, to the conditional density, the 6th order approximation, $p^6(\cdot \mid X_0, X_\delta^I)$, and 10^4 times their difference.

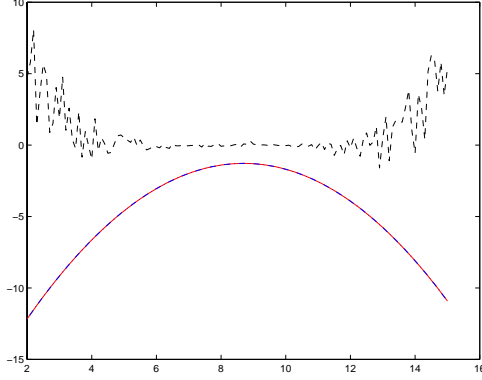


Figure 14. The first-order approximation, $\log(p^1(\cdot \mid X_0, X_\delta^I))$, to the log-conditional density, the 6th order approximation, $\log(p^6(\cdot \mid X_0, X_\delta^I))$, and 10^4 times their difference.

i	$ p^i - p^6 $	$\frac{ p^i - p^6 }{p^6}$	$ \log(p^i) - \log(p^6) $	$\frac{ \log(p^i) - \log(p^6) }{ \log(p^6) }$
1	0.0067	1.0492	1.0493	0.1405
2	0.0065	0.6289	0.6289	0.0990
3	0.0060	0.5913	0.5913	0.0918

Table III: The first column is 10^4 times the average absolute differences between the density approximation, p^i , of order i and that of order 6. The second column shows 10^4 times the average absolute relative differences. The third and fourth columns present the same information for the log-densities (also multiplied by 10^4).

Appendix C: Risk-Free Parameters

The estimates of the parameters for the reference model are reported in Table IV. We note that all²² parameters are significant at the 5% level. We also note that the standard deviations computed using the first and second derivatives are similar, indicating that the asymptotic analysis may be appropriate.

	est.	SD1	SD2
K^{vv}	0.0047	(0.0030)	(0.0021)
K^{rv}	-0.0268	(0.0078)	(0.0109)
K^{rr}	0.3384	(0.1378)	(0.1548)
θ^v	107.4039	(—)	(—)
θ^r	5.6800	(—)	(—)
Σ^{rv}	0.0436	(0.0126)	(0.0145)
Σ^{rr}	0.1145	(0.0353)	(0.0249)
λ^v	0.0000	(—)	(—)
λ^r	-0.0764	(0.0372)	(0.0412)

Table IV: The first column contains the estimates. The second and third columns contains standard deviations estimated using first (SD1) and second (SD2) derivatives, respectively.

References

- Ait-Sahalia, Yacine, 1999, Maximum Likelihood Estimation of Discretely Sampled Diffusions: A Closed-Form Approach, Princeton University.
- Brandt, Michael W., and Pedro Santa-Clara, 1999, Simulated Likelihood Estimation of Multivariate Diffusions with an Application to Interest Rates and Exchange Rates with Stochastic Volatility, Working Paper.
- Bulow, Jeremy, and Kenneth Rogoff, 1989a, A Constant Recontracting Model of Sovereign Debt, *Journal of Political Economy* 97, 155–78.
- Bulow, Jeremy, and Kenneth Rogoff, 1989b, Sovereign Debt: Is to Forgive to Forget?, *American Economic Review* 79, 43–50.
- Chen, R.-R., and L. Scott, 1993, Maximum Likelihood Estimation for a Multifactor Equilibrium Model of the Term Structure of Interest Rates, *Journal of Fixed Income* 3, 14–31.
- Collin-Dufresne, Pierre, and Bruno Solnik, 2001, On The Term Structure of Default Premia in the Swap and LIBOR Markets, *Journal of Finance* 56, 1095–1116.
- Cox, John C., Jonathan E. Ingersoll, and Stephen A. Ross, 1985, A Theory of the Term Structure of Interest Rates, *Econometrica* 53, 385–407.

- Dai, Qiang, and Kenneth J. Singleton, 2000, Specification Analysis of Affine Term Structure Models, *Journal of Finance*, 55, 1943–1978.
- Dai, Qiang, and Kenneth J. Singleton, 2001, Expectation Puzzles, Time-Varying Risk Premia, and Dynamic Models of the Term Structure, forthcoming, *Journal of Financial Economics*.
- Davidson, Russell, and James G. MacKinnon, 1993, *Estimation and Inference in Econometrics*. (Oxford University Press Oxford).
- Delbaen, F., and W. Schachermayer, 1999, A General Version of the Fundamental Theorem of Asset Pricing, *Mathematische Annalen* 300, 463–520.
- Dellacherie, C., and P.-A. Meyer, 1978, *Probabilities and Potential*. (North-Holland Amsterdam).
- Deutsche Bank Research, 1998a, Emerging Markets Debt Strategy, Deutsche Bank.
- Deutsche Bank Research, 1998b, Russian Debt Restructuring: Global Implications, (August 26).
- Duffee, Gregory R., 1999, Estimating the Price of Default Risk, *Review of Financial Studies* 12, 197–226.
- Duffee, Gregory R., 2000, Term Premia and Interest Rate Forecasts in Affine Models, Haas School of Business, Univ. California, Berkeley, forthcoming, *Journal of Finance*.
- Duffee, Gregory R., and Richard Stanton, 2001, Estimation of Dynamic Term Structure Models, Haas School of Business, Univ. California, Berkeley.
- Duffie, Darrell, 1998, First-to-Default Valuation, Graduate School of Business, Stanford University.
- Duffie, Darrell, and Ming Huang, 1996, Swap Rates and Credit Quality, *Journal of Finance* 51, 921–950.
- Duffie, D., and R. Kan, 1996, A Yield-Factor Model of Interest Rates, *Mathematical Finance* 6, 379–406.
- Duffie, D., J. Pan, and K. Singleton, 2000, Transform Analysis and Asset Pricing for Affine Jump Diffusions, *Econometrica* 68, 1343–1376.
- Duffie, Darrell, and Kenneth J. Singleton, 1997, An Econometric Model of the Term Structure of Interest Rate Swap Yields, *Journal of Finance* 52, 1287–1323.

- Duffie, Darrell, and Kenneth J. Singleton, 1999, Modeling Term Structures of Defaultable Bonds, *Review of Financial Studies* 12, 687–720.
- Dullmann, K., and M. Windfuhr, 2000, Credit Spreads Between German and Italian Sovereign Bonds– Do Affine Models Work?, Working Paper, University of Mannheim.
- Eaton, Jonathan, and Mark Gersovitz, 1981, Debt with Potential Repudiation: Theoretical and Empirical Analysis, *Review of Economic Studies* 48, 289–309.
- Eichengreen, Barry, and Ashoka Mody, 2000, Would Collective Action Clauses Raise Borrowing Costs?, NBER Working Paper No. 7458.
- Fischer, E., R. Heinkel, and J. Zechner, 1989, Dynamic Capital Structure Choice: Theory and Tests, *Journal of Finance* 44, 19–40.
- Fisher, M., and C. Gilles, 1996, Estimating Exponential Affine Models of the Term Structure, Working paper, Federal Reserve Board, Washington DC.
- Gallant, A. R., and G. Tauchen, 1996, Which Moments to Match?, *Econometric Theory* 12, 657–681.
- Gibson, Rajna, and Suresh M. Sundaresan, 1999, A Model of Sovereign Borrowing and Sovereign Yield Spreads, Graduate School of Business, Columbia University.
- Harrison, M., and D. Kreps, 1979, Martingales and Arbitrage in Multiperiod Securities Markets, *Journal of Economic Theory* 20, 381–408.
- IMF, 1998, World Economic Outlook: Interim Assessment, World Economic and Financial Surveys (Washington).
- IMF, 1999a, International Capital Markets, World Economic and Financial Surveys (Washington).
- IMF, 1999b, Russian Federation: Recent Economic Developments, IMF Staff Country Report No. 99/100.
- IMF, 2000a, International Capital Markets, International Monetary Fund.
- IMF, 2000b, Russian Federation: Selected Issues, IMF Staff Country Report No. 00/150.
- Jarrow, R., and S. Turnbull, 1995, Pricing Derivatives on Financial Securities Subject to Credit Risk, *Journal of Finance* 50, 53–85.

- Keswani, Aneel, 1999, Estimating a Risky Term Structure of Brady Bonds, London Business School.
- Kremer, Michael, and Paras Mehta, 2000, Globalization and International Public Finance, Working Paper, Harvard University.
- Lando, David, 1998, On Cox Processes and Credit Risky Securities, *Review of Derivatives Research* 2, 99–120.
- Liu, Jun, Jun Pan, and Lasse Heje Pedersen, 1997, Density-Based Inference in Affine Jump-Diffusions, Graduate School of Business, Stanford University.
- Longstaff, F., and E. Schwartz, 1995, A Simple Approach to Valuing Risky Fixed and Floating Rate Debt, *Journal of Finance* 50, 789–819.
- Malleret, Thierry, Natalia Orlova, and Vladimir Romanov, 1999, What Loaded and Triggered the Russian Crisis?, *Post-Soviet Affairs* 15, 107–129.
- Merrick, John J., 1999, Crisis Dynamics of Implied Default Recovery Ratios: Evidence From Russia and Argentina, Stern School of Business, New York University.
- Merton, R., 1974, On The Pricing of Corporate Debt: The Risk Structure of Interest Rates, *The Journal of Finance* 29, 449–470.
- Pagès, Henri, 2000, Estimating Brazilian Sovereign Risk from Brady Bond Prices, Working Paper, Bank of France.
- Pearson, N. D., and T. Sun, 1994, Exploiting the Conditional Density in Estimating the Term Structure: An Application to the Cox, Ingersoll, and Ross Model, *Journal of Finance* 43, 1279–1304.
- Pedersen, Asger Roer, 1995, A New Approach to Maximum Likelihood Estimation for Stochastic Differential Equations Based on Discrete Observations, *Scandinavian Journal of Statistics* 22, 55–71.
- Protter, P., 1990, *Stochastic Integration and Differential Equations*. (Springer-Verlag New York).
- Schönbucher, Philipp J., 1998, Term Structure Modelling of Defaultable Bonds, *Review of Derivatives Research* 2, 161–192.

Notes

¹GKO's are Ruble-denominated, short-term discount bonds and OFZs are longer-dated coupon bonds issued by the Russian Treasury. In mid 1998, Russia's Treasury debt had reached \$70 billion, of which about one third was held by nonresident investors (IMF (1998)).

²Additive errors were assumed in the studies by Chen and Scott (1993), Duffie and Singleton (1997), and Duffie (1999), among others. In its most general form, our modeling framework is agnostic about the determinants of the bond-specific factors. In circumstances for which a researcher has data on issuer-specific or macro-economic variables that are useful for distinguishing among various credit and liquidity components of sovereign spreads, then it may be possible to identify bond-specific components for each bond rather, as we do here, focus on spreads relative to a benchmark bond.

³Liu, Pan, and Pedersen (1997) propose a closed-form approximation to the likelihood function of a general affine jump-diffusion, Ait-Sahalia (1999) offers closed-form approximations to the likelihood function of a generic one-dimensional diffusion, and Pedersen (1995) derives a simulation-based approximation to the likelihood function of a generic diffusion (see also Brandt and Santa-Clara (1999)).

⁴The assumption that the volatility of each component of X_t^I is an affine transformation of the square-root of itself is essentially without loss of generality in affine diffusion models. As will be made precise subsequently, it is the independence assumption that leads us to lose generality by ruling out feedback among the components of X_t^I through their drifts.

⁵An exception is Keswani (1999), who applies the model of Longstaff and Schwartz (1995) to Brady bonds.

⁶The short-rate process, r , is a predictable process with $\int_0^t |r_s| ds < \infty$ a.s., and $E^Q \left[\exp(-\int_0^t r_s ds) \right] < \infty$ for all t . See Protter (1990) for definitions not provided here.

⁷An intensity process h for N is defined by the fact that h is non-negative and predictable, and by the key property that $N_t - \int_0^t h_s ds$ defines a mar-

tingale under Q .

⁸Here, Y_n and W_n are random variables with outcomes in $[0, 1]$ that are measurable with respect to $\mathcal{F}_{\max(\tau_n, T)}$, where $\tau_n = \inf\{t : N_t \geq n\}$.

⁹Specifically, L is a predictable process such that for each $n \in \mathbb{N}$, $L_\tau = E^Q(1 - Y_n - W_n \mid \mathcal{F}_{\tau-})$ where τ is the stopping time $\tau = \inf\{t : N_t \geq n\}$. Dellacherie and Meyer (1978) show that such a process exists. (See Duffie (1998) for further details on this construction in a default setting.)

¹⁰That is, $\Delta V_t \Delta N_t = 0$ almost surely for all t . From its definition, V cannot ever jump if R and Z are measurable with respect to an observable diffusion process, as in our later parametric model. We could relax this and obtain our pricing formula with a doubly-stochastic model, under which the counting process N is, conditional on a sub-filtration to which h is adapted, Poisson.

¹¹That is, we specify that, under Q , $\tilde{W}_t^{v,r}$ is a standard Brownian motion in \mathfrak{R}^2 .

¹²See, for example, Chen and Scott (1993), Pearson and Sun (1994), and Duffie and Singleton (1997). In affine models, the prices of zero-coupon bonds are exponential-affine functions of the state variable (Duffie and Kan (1996)) with easily computed coefficients. Further, the yields on coupon bonds are simple functions of the zero-coupon prices, so g is known in closed-form up to the solution of certain ordinary differential equations underlying the zero-coupon pricing.

¹³That is, we first fix θ^r at the historical mean of the one-year treasury yield between 2/1/1962 and 7/9/1999 (which is 6.82%), then estimate the model and compute the sample mean of the implied short rate to be 5.68%, and finally we fix θ^r at this level and re-estimate. The estimates obtained in the last stage are very similar to those from the preceding stage, and the sample mean of the implied short rate remains 5.68%.

¹⁴This representation (11) is equivalent to assuming that s is an affine function of a trivariate state vector (r, v, η) that is an affine diffusion with (possibly) correlated components. Additional state variables that drive s , but not r directly, can be accommodated. Given, however, our limited sample size and the fact that we are allowing for each bond spread to have an

idiosyncratic component, we proceed with this parsimonious representation.

¹⁵More precisely, using Bayes' Rule, $p(c, b; \psi, \phi)$ can be expressed as

$$p(c; \psi) \prod_t p(s_t, \gamma_t \mid v_{t-1}, r_{t-1}, s_{t-1}, \gamma_{t-1}, v_t, r_t; \psi, \phi) \frac{1}{|\det D_{s,\gamma} H|},$$

so we estimate ϕ by maximizing the partial likelihood function

$$\hat{\phi} = \arg \max_{\phi} \prod_t p(s_t, \gamma_t \mid v_{t-1}^{\hat{\psi}}, r_{t-1}^{\hat{\psi}}, s_{t-1}, \gamma_{t-1}, v_t^{\hat{\psi}}, r_t^{\hat{\psi}}; \hat{\psi}, \phi) \frac{1}{|\det D_{s,\gamma} H|}.$$

¹⁶ It can be seen that the stationary distribution of Z is Beta, with form parameters $2\kappa^z \theta^z / (\sigma^z)^2$ and $2\kappa^z (1 - \theta^z) / (\sigma^z)^2$. At the time, τ , of a credit event, we can suppose that Z_τ is drawn from this stationary distribution.

¹⁷Letting $X_t' = (r_t, v_t, s_t, N_t, Z_t)$ denote the state vector and $X_t^{(1)'} \equiv (r_t, v_t, s_t, N_t)$, the likelihood of the data, over the pre- and post-default periods, is

$$\begin{aligned} p(c, b, N; \hat{\psi}, \phi, \zeta) &= p(c; \hat{\psi}) \\ &\cdot \left[\prod_{t=1}^{\tau-1} p(s_t \mid X_{t-1}^{(1)}, v_t, r_t; \phi) p(N_t \mid N_{t-1}, s_{t-1}, s_t) p(\gamma_t \mid \gamma_{t-1}; \phi) \frac{1}{|\det DH_s(X_t^{(1)})|} \right] \\ &\cdot \left[\prod_{t=\tau}^T p(s_t \mid X_{t-1}^{(1)}, v_t, r_t; \phi) p(N_t \mid N_{t-1}, s_{t-1}, s_t) p(Z_t \mid X_{t-1}, X_t^{(1)}; \zeta) \right] \\ &\cdot \left[\frac{1}{|\det DH_{s,Z}(X_t)|} \right], \end{aligned}$$

where τ is the time of the credit event, γ is (as before) the vector of bond-specific spread factors for the pre-default period, and $H(r_t^\psi, v_t^\psi, s_t, N_t, Z_t; \phi)$ is the model counterpart to the observed price vector b_t .

¹⁸By partial likelihood function we mean the portion of the log-likelihood function that is relevant for estimation of ϕ . *Independent Recovery* under the actual measure allows us to write $p(Z_t \mid X_{t-1}, X_t^{(1)}; \zeta)$ as $p(Z_t \mid Z_{t-1}; \zeta)$.

Though, in this case, it would be reasonable to assume that η and ϕ are not functionally related, this strong independence condition is not sufficient to interpret (20) as the partial likelihood function. The reason is that the state variable Z^ϕ inverted from the pricing model would depend indirectly on ϕ and, hence, so would $p(Z_t^\phi, | Z_{t-1}^\phi; \zeta)$.

¹⁹ The model prices of the out-of-sample bonds are computed as follows. First, we take the write down of each out-of-sample bond to be equal to the write down of the benchmark bond (MinFin 4). Second, we take the short spread of an out-of-sample bond to be a convex combination of the short spreads of the two in-sample bonds with the nearest maturity. The weight of each in-sample bond in the convex combination depends linearly on the closeness in maturity.

²⁰Under regularity, most of the affine diffusions studied in the literature on dynamic term structure models can be transformed to this form. See Dai and Singleton (2000).

²¹When $\alpha^D > 0$, V_δ^J is bounded away from zero and dominated convergence is immediate. Otherwise, it may be shown under regularity.

²²The parameter K^{vv} is significant when we use the standard error computed using the second derivative, but not with that computed using the first derivative.