ESSAYS ON DERIVATIVES

Rodolfo Alejandro Oviedo-Helfenberger
Faculty of Management
McGill University, Montreal

May 2005

The Suboptimality of Early Exercise of Futures-Style Options: A Model-Free Result, Robust to Market Imperfections and Performance Bond Requirements

Improving the Design of Treasury-Bond Futures Contracts Contents

The Determinants of Credit Default Swap Premia

The Suboptimality of Early Exercise of Futures-Style Options: A Model-Free Result, Robust to Market Imperfections and Performance Bond Requirements

Rodolfo Oviedo*
Faculty of Management, McGill University
Facultad de Ciencias Empresariales, Universidad Austral

First Version: January 22, 2005
This Revision: December 22, 2006

Abstract

Options on futures traded in Eurex, Euonext-LIFFE and SFE are subject to futures-style premium posting (FSPP): the premium is not paid up front but marked to market. In the U.S., steps to adopt FSPP have been taken. The previous literature has shown that early exercise of such options is neither optimal for Black's dynamics of the underlying futures price, nor in an equilibrium CIR setting. The paper proves that these options, either calls or puts, are never optimally exercised early under any possible dynamics of the underlying and interest rates. It also proves that, unlike the case of plain-vanilla calls on non-dividend-paying stock, the previous result is robust to transaction costs, illiquidity and collateral requirements. Hence, assuming that investors are rational, it is superfluous to keep the American feature among the specifications of futures-style options.

JEL Classification: G13

Keywords: futures-style; early exercise; options on futures; transaction costs; performance bond; SPAN; put-call parity, LIFFE, EUREX, SFE.

*The author holds the ROFEX Chair of Derivatives at Universidad Austral. He is grateful to the Institut de Finance Mathématique de Montréal (IFM^2) for generous financial support. He would like to thank Marie-Claude Beaulieu, Laurette Dubé, Jan Ericsson, Pascal François, Ruth Hanna, Kris Jacobs and Stefano Mazzotta for helpful comments. Any remaining inadequacies are the author's alone. Correspondence to: Rodolfo Oviedo, Facultad de Ciencias Empresariales, Universidad Austral, Paraguay 1950, Rosario, SF, Argentina S2000FZF; Tel: +54(341) 522-3083; Fax: +54(341) 522-3001; E-mail: rodolfo.oviedo@fce.austral.edu.ar

1 Introduction

The premium of traditional options on futures is paid up front. It is well known that such options, both calls and puts, are optimally exercised before expiration, should they be deep enough in the money. Using a model-free framework, this paper shows that exercising options on futures, either calls or puts, before expiration is suboptimal when the premium is not paid up front but marked to market in the same way as the price of a futures contract.

Options can be classified according to their underlying into options on spot and options on futures. While the former give the long the right to buy or sell spot, the latter give the right to buy or sell futures. In addition, options can be qualified as either traditional- or futures-style as per the method of posting the premium. On the one hand, the premium of a traditional-style option is paid up front. This way of paying the premium is referred to as "traditional-style premium posting". On the other hand, the premium of a futures-style option is settled in the same way as the price of a futures contract. The buyer of a futures-style option does not pay the full option premium at the initiation of the transaction. Rather, during the life of the option (including the exercise day), he pays (receives) any decrease (increase) in the option's premium as a margin. For that purpose, the exchange determines a settlement premium according to market conditions at the end of the trading hours of each trading day. Upon exercise, the buyer pays the settlement premium of the day. The settlement premium paid upon exercise plus the net cumulative margin paid during the life of the option add up to the premium originally negotiated. Under the method just described, the premium of the option is said to be futures-style posted.

At present, all exchange-traded options on spot are traditional-style. However, options on futures can be traditional-style or futures-style, depending on the exchange where they are listed. This paper deals with *futures-style options on futures*.

For parsimony, the literature uses the abbreviated terminology initiated by Duffie (1989): *pure futures options* (PFO).[1] All options on futures traded on EUREX, Euronext and SFE are PFO.[2] Euronext is a holding company comprising the operations of AEX, Belfox, BVLP, BXS, LIFFE, MATIF and MONEP. Accepting a petition by American exchanges, the Commodity Trading Futures Commission has permitted futures-style premium posting for commodity options on futures.

Futures-style options offer advantages over traditional-style options: i) With traditional-style premium posting, the writer has to maintain collateral, known as the *premium margin*, at the exchange's disposal for an amount at all times equal to the last settlement premium.[3] This collateral is kept by the exchange and not passed to the buyer; therefore, there is an aggregate commitment of liquidity by the traders.[4] This commitment can be substantial when an option is in the money. Futures-style options avoid any such immobilization of collateral. A premium margin is not needed because, being settled as futures, these options have zero net present value and, therefore, the exchange can costlessly replace the futures-style position of an insolvent party.[5] ii) The price of a traditional-style, in-the-money option can be discouraging for a potential buyer with capital restrictions. In contrast, no down payment is required with futures-style options. iii) Although usually considered unlikely, a clearing house default can hurt a traditional-style more than a futures-style option buyer. While the buyer of traditional-style option has an exposure at all times equal

[1]Futures-style options on spot are studied by Kutner, Porter & Thatcher (2001) who, adapting Barone-Adesi & Whaley (1987), show that an early exercise premium can exist when the cost of carry is negative.

[2]For an example of contract specifications, see Eurex (2004).

[3]If the short becomes insolvent, the exchange will use the premium margin to buy an option of the same strike and expiration. The exchange's counterparty will receive a short position in the option and thus replace the insolvent trader.

[4]An economic loss may derive if the return of the collateral is lower than the opportunity cost of the trader.

[5]Exchanges also require collateral to cover losses from potential changes in prices. I do not expand on this extra requirement here because it is essentially the same for both traditional-style and futures-style options. Later, I will return to this topic when dealing with the optimality of early exercise.

to the option value, the purchaser of futures-style options is, upon resettlement, risk free.

The literature on futures-style options is limited.[6] Two papers show the suboptimality of early exercise of PFO under particular models. Lieu (1990) proves suboptimality assuming a Black (1976) market, where the futures price follows a geometric Brownian motion and the interest rate is constant. Chen & Scott (1993) prove the result under stochastic interest rates within the Cox, Ingersoll & Ross (1985a) general equilibrium framework and rely on Cox, Ingersoll & Ross (1981) to establish a link between the futures price and the spot price of the underlying. Both papers demonstrate that, within their respective settings, the European PFO price exceeds its intrinsic value and, therefore, so does the price of the American counterpart, thus making early exercise suboptimal.[7] However, after deriving a model for European PFO, where the futures price and the interest rate are driven by the same source of uncertainty, Kuo (1993) posits that early exercise might be priced in his model, leaving the valuation of American PFO for future research.

The most important contribution of this paper is to prove the suboptimality of early exercise of PFO in a model-free context. I do so in the spirit of Merton (1973), assuming only that investors prefer more to less. Specifically, the stochastic process of the underlying futures can be absolutely arbitrary, provided there exists a positive probability that the option will expire either in or out of the money. Moreover, I make no assumptions about the dynamics of interest rates. For example, futures price and interest rates dynamics can display stochastic volatility, jumps in the level and volatility, and discontinuities in time. Furthermore, the proofs do not rely on the existence of either an interest rate market or a spot market for the underlying of

[6] Research dealing with PFO but not specifically with the early exercise strategy include Duffie & Stanton (1992), Twite (1996), Satchell, Stapleton & Subrahmanyam (1997), and Frey & Sommer (1998), and White (1999).

[7] Empirical work by Kutner et al. (2001) shows that the early exercise premium of pure futures calls traded on the Australian All Ordinaries Share Price Index is economically zero.

the futures contract. Apart from their generality, the arguments used in this paper have two other advantages: first, their simplicity and, second, the analogy to the standard arguments used for options on spot by the literature.

One may wonder, at this point, why this line of argument has not been tried before by the literature when it would only have been natural to follow Merton (1973) and regular textbooks in their treatment of options on spot, where, before studying specific models, they derive model-free rational constraints to options premia and exercise policy. The reason for this omission can be attributed to the particularities of PFO: (i) both the option and its underlying always have zero value, and (ii) there are cash flows during the whole life of the option. Therefore, we cannot structure arbitrage or dominance arguments in the usual way, whereby only two points in time are considered. Hence, it may seem that a model for European options is needed to show that the premium always surpasses the intrinsic value, and that, therefore, the dominating American counterpart should not be exercised early.

The scope of this paper extends beyond merely allowing for arbitrary futures price or interest rate dynamics in a frictionless market. Early exercise is still suboptimal in a market with transaction costs, liquidity constraints and collateral requirements. This ensures the realism of the conclusions, from which I derive implications for the design of PFO contracts. All the results of this paper hold regardless of the underlying of the futures contract: stocks, bonds, commodities with convenience yields, indexes representing traded assets (e.g., stock indexes), or indexes representing non-traded quantities (e.g., weather indexes).

Lieu (1990) and Chen & Scott (1993) provide proofs of put-call parity in their respective settings.[8] As previously mentioned, they also prove that, before expiration,

[8] Easton (1997) tests this put-call parity using data of four major contracts traded on the Sydney Futures Exchange and finds that "The precise parity relationship was observed in between 15% and one third of all cases, depending on the contract. The only systematic violation detected is that in-the-money put and call options are found to be underpriced by a small amount when compared with the parity relationship."

the premium of a PFO always exceeds its intrinsic value. In addition, Lieu (1990) shows that the intrinsic value of calls and puts with the same strike and expiration have equal time value. I prove that, in frictionless markets, these three results are totally general.

All results by Lieu (1990) and Chen & Scott (1993) assume a frictionless market. In this setting, I use arbitrage arguments and backward induction to prove them. In the case of markets with frictions, the basic tool for the proofs is the gains process (cumulative marking-to-market proceeds) of PFO, of their underlying futures, and of portfolios combining both instruments. I introduce a novel dominance criterion based on gains processes of alternative strategies. Under this criterion, the early-exercise strategy is dominated by the holding strategy, even in the presence of transaction costs, liquidity constraints, and collateral requirements. It is noteworthy that this degree of generality and realism does not hold in other settings. For example, it is well known that a traditional-style option on a non-dividend paying stock is never optimally exercised early under the usual assumption of a frictionless market. However, such option may be optimally exercised early in a more realistic setting, as the paper exemplifies.

The reminder of the paper is organized as follows. The features of PFO contracts are described in Section 2. Put-call parity, the equality of time values of calls and puts with the same strike and expiration, their positivity, and the suboptimality of early exercise are studied in Section 3 assuming a frictionless market. The suboptimality of early exercise is robust to the introduction of transaction costs and liquidity constraints in Section 4, and of collateral requirements in Section 5. Section 6 draws some implications for the design of option contracts and Section 7 concludes.

2 Futures-style options

I begin by describing regular futures with physical delivery. My objective is to start with a well-known instrument and then draw an analogy with PFO.

Let F_0 be a futures price agreed upon between the buyer and the seller *during the trading hours*, and $F_1,\ F_2, ...,\ F_{T^f}$ the *settlement prices* that the exchange determines according to the price prevalent at closing time each day. The subindex 1 corresponds to the day of initiation of the futures contract; the subindex 2, to the following day, and so on, and T^f to the expiration. The supraindex f of T is meant to distinguish the expiration T^f of the futures contract from that of an option on this underlying, denoted T. Note that the subindex $t = 0, 1, 2, ..., F_{T^f}$ of F_t does not represent a cardinal number: both indexes 0 and 1 correspond to the day of trade: 0 to the time, during trading hours, when the trade occurs, and 1, to the closing time.

The cash flows for the buyer are:

Day	**Cash flows**
1 (Transaction day)	$F_1 - F_0$
2	$F_2 - F_1$
...	...
$T^f - 1$	$F_{T^f-1} - F_{T^f-2}$
T^f (Expiration)	$F_{T^f} - F_{T^f-1}$

Each of these cash flows is referred to as a variation margin and, the collection of them, as marking-to-market proceeds. At expiration, the long pays an invoice price equal to F_{T^f}, and receives the underlying. In summary, the total cash flows for the buyer can be grouped into three components:[9]

[9] In this section, I will neglect any return or financing cost derived from the variation margins.

A1: $F_{T^f} - F_0$, which is the sum of the cash flows of the previous table, i.e., the cumulative variation margin at expiration.

A2: $-F_{T^f}$, the delivery price, and

A3: S_{T^f}, the market value of the underlying received.

Note that A1 plus A2 add up to a net payment of F_0, the original futures price agreed upon. This is paid in exchange for the underlying, whose price turns out to be S_{T^f} at expiration. Therefore, the final profit is $S_{T^f} - F_0$.[10]

A call (put) option contract on futures gives the long the right but not the obligation to buy (sell) a futures contract at a prespecified exercise price K up to a certain date $T \leq T^f$. The variation margin of the futures contract created upon exercise of a call is $F_\tau - K$, where τ is an index indicating the closing time of the exercise date, and F_τ is the corresponding settlement price. An equivalent definition of a call on futures is: an option whose exercise generates

- a payoff $F_\tau - K$ that is cash settled at time τ, and
- a long futures contract initiated at F_τ (instead of K).

From this point of view, which I will use in this paper, $F_\tau - K$ is not a futures margin but an *exercise cash flow*. The futures contract originated has zero net present value at τ, if the settlement price accurately represents the futures price at that time. In the case of a put, the long receives an exercise cash flow of $K - F_\tau$ and a short futures contract initiated at F_τ.

If an option contract is settled futures-style, the premium is not paid cash; instead, any increase (decrease) of the premium generates a positive (negative) variation margin to the long, and a negative (positive) variation margin to the short. Thus, a

[10] In the absence of frictions, no-arbitrage requires that $F_T = S_T$; therefore, A2 and A3 add up to zero and the final profit is given by A1.

futures-style option works exactly as a futures contract while the option is alive, i.e., while neither exercise time τ nor expiration T has taken place. If, for example, the last settlement premium is zero, the long will have paid the entire premium during the life of the option.

For concreteness, I will consider a call option. (For the case of a put, it is enough to replace C by P, and $F_\tau - K$ by $K - F_\tau$.) Assuming the buyer exercises the call, the variation margins are:

Day	Cash flow
1 (Transaction day)	$C_1 - C_0$
2	$C_2 - C_1$
...	...
$\tau - 1$	$C_{\tau-1} - C_{\tau-2}$
τ (Exercise day)	$C_\tau \quad - C_{\tau-1}$

where C_0 is the premium negotiated during trading hours, C_1 is the settlement premium of the day of trade, C_2 the settlement premium of the next day, and so on. On the day the option is exercised, the long receives an exercise cash flow of $F_\tau - K$, and pays the call's settlement price of that day, C_τ. Summing up, the total cash flows for the buyer can be arranged into three components:

B1: $C_\tau - C_0$, the accumulated marking-to-market proceeds,

B2: $-C_\tau$, the settlement price of the day of exercise, and

B3: $F_\tau - K$, the exercise cash flow.

The analogy with a futures contract with physical delivery is that $F_\tau - K$ is like the commodity received and C_τ is the invoice price paid for the underlying; however, in the case of a PFO, $F_\tau - K$ is received in cash. The sum of B2 and B3, $F_\tau - K - C_\tau$, will be referred to as *net exercise cash flow*.

B1 plus B2 add up to $-C_0$, where C_0 is the premium originally negotiated. Adding B3, the exercise cash flow, gives the net payoff

$$(F_\tau - K) - C_0, \tag{1}$$

which is the same as that of a traditional-style option on futures, if we abstract from the timing of the cash flows.

When the option is not exercised, τ is replaced by T in B1, B2, and the above table; and B3 disappears. When the option expires out of the money, the exchange sets $C_T = 0$; therefore, B2 is null, and B1 is $-C_0$.

Example 1

Call expiring out of the money

			K	100	Variation Margin	Acc. Var. Margin
Day 1	F_0	101	C_0	1.39		
Day 1	F_1	100	C_1	0.74	-0.65	-0.65
Day 2	F_2	100	C_2	0.52	-0.22	-0.86
Day 3	F_3	99	C_3	0	-0.52	-1.39
Day 4	F_4	98				

This table considers a pure futures call with strike 100 that is purchased on Day 1 and expires on Day 3. The last two columns present cash flows from the buyer's perspective. At the moment of transaction, the premium was $C_0 = 1.39$ (and the futures price was 101). Had the settlement price of Day 1 been $C_1 = 1.39$, no payment would have been necessary for that day. However, a lower settlement premium,

$C_1 = 0.74$, generated a negative cash flow, $0.74 - 1.39 = -0.65$. Note that the last settlement premium is $C_3 = 0$ because the option expires out of the money. The cumulative variation margin is -1.39. The agent has lost all the premium.

Example 2

Call expiring in the money

			K	100	Variation Margin	Acc. Var. Margin
Day 1	F_0	100	C_0	0.78		
Day 1	F_1	101	C_1	1.35	0.57	0.57
Day 2	F_2	102	C_2	2.04	0.69	1.26
Day 3	F_3	101	C_3	1	-1.04	0.22
Day 4	F_4	100				

In the example of this table, the buyer not only does not pay the premium $C_0 = 0.78$ up front, but he is credited a margin of 0.57 on the very transaction day.

If the long had decided to exercise on Day 2, he would have received an exercise cash flow of $102 - 100 = 2$. That same day, he would have paid the last settlement premium $C_2 = 2.04$, resulting in a *net* exercise cash flow of $2 - 2.04 = -0.04$. Adding the cumulative variation margin 1.26, would have resulted in a profit to the buyer: $1.26 - 0.04 = 1.22$.[11] The contracted premium would have been paid as the sum of the accumulated variation margins minus the last settlement price paid upon exercise: $1.26 - 2.04 = -0.78$. As with traditional-style options, the net profit would have been equal to the exercise cash flow minus the premium: $2 - 0.78 = 1.22$.

At expiration (Day 3), the exchange fixes the last settlement price C_3 equal to exercise cash flow $101 - 100 = 1$; hence, the *net* exercise cash flow is $1 - 1 = 0$.

[11] Although the long would have made a profit of 1.22, exercising would have prevented him from making a profit of 1.26, hence rendering early exercise suboptimal. The difference is the negative net exercise cash flow -0.04.

Whenever this is the case, the profit is equal to the cumulative variation margin, in this case 0.22. As always, the cumulative variation margin minus the last settlement premium yields the contracted premium $0.22 - 1 = -0.78$. The exercise cash flow minus this premium gives the profit $1 - 0.78 = 0.22$.

3 Case of a frictionless market

The sketch of this section is the following. First, I prove put-call parity, which leads to the equality of time values of calls and puts with the same strike and expiration. Building on this result and on the positivity of PFO's premia, I demonstrate the positivity of time values before expiration, whence the suboptimality of early exercise follows. A frictionless market is assumed throughout.

Assumption 1 *There are no arbitrage opportunities.*

First, note that the intrinsic values trivially satisfy put-call parity:[12]

$$(F_t - K)^+ - (K - F_t)^+ \equiv F_t - K. \tag{2}$$

For $t = T$, we have put-call parity for options prices at expiration, $C_T - P_T = F_T - K$. Then, following a recursive no-arbitrage argument from expiration to the valuation time, and going through all intermediate settlement times, obtains the European put-call parity:

Theorem 1 *Put-call parity for European pure futures options is*

$$c_t - p_t = F_t - K, \tag{3}$$

[12]Notation: i) $(X)^+ = \max\{X, 0\}$. ii) Recall that $t = 0, 1, 2, ..., T$. The subindex 0 denotes any time during trading hours of the day of trade; the subindex 1, the closing time of that day; the subindex 2, the closing time of the following day, and so on. The case $t = 0$ makes the the following expressions completely general in terms of time.

where c_t and p_t stand for the premia of European calls and puts with the same strike and expiration.

The proof of this theorem, given in the Appendix, extends the validity of Lieu (1990)'s put-call parity by only assuming the absence of arbitrage. While he proves this result on the basis of a strategy whereby the variation margins are reinvested at a constant risk-free rate, I use a no-arbitrage backward induction argument wherein the arbitrageurs set up one-day zero-investment portfolios (comprising futures-style contracts), thereby rendering interest rates irrelevant.

Theorem 2 *European pure futures puts and calls with the same expiration and strike have equal time values:*

$$c_t - (F_t - K)^+ = p_t - (K - F_t)^+ \,. \tag{4}$$

Proof. This result can be obtained by subtracting the put-call parity (2) of the intrinsic values from the put-call parity (3) of the option prices, and rearranging terms. ■

Assumption 2 *Before expiration, there is always a positive probability that an option will expire in the money and a positive probability that it will expire out of the money. More formally, for any time t such that $t < T$, we have $\Pr_t(F_T > K) > 0$ and $\Pr_t(F_T < K) > 0$.*

Lemma 1 *PFO's premia are always positive before expiration.*[13]

This lemma is proved in the Appendix by using Assumption 2 at $t = T - 1$ and then following a recursive arbitrage argument.

[13] For a traditional-style option, we can argue that it gives a right but not an obligation and, therefore, will always generate a positive cash flow on exercise. This argument does not apply to futures-style options because the buyer of the option is obliged to pay variation margins whenever the settlement premium drops.

By utilizing this lemma and the equality of the time values of puts and calls with the same strike and expiration, the next theorem is easy to demonstrate.

Theorem 3 *Before expiration, the time value of a PFO is always positive:*

$$c_t > (F_t - K)^+ \qquad\qquad p_t > (K - F_t)^+ .$$

Proof. (a) For an at- or out-of-the-money option, the theorem follows from Lemma 1. (b) To prove the theorem for an in-the-money call, consider that a put with the same strike and expiration is out of the money and, by (a), has a positive time value. From this and Theorem 2, it follows that the time value of an in-the-money call is positive. A symmetric argument establishes the positivity of the time value of an in-the-money put. ■

Theorem 4 *It is never optimal to exercise a PFO before expiration.*

Proof. Let τ be the settlement time of a potential early-exercise date. Because $C_t \geq c_t$, Theorem 3 implies that $C_t > F_t - K$ for $t < T$. Therefore, the net exercise cash flow of a call is negative prior to expiration, $F_t - K - C_t < 0$. Since the generated futures contract has zero net present value, that negative cash flow is sufficient to rule out optimal early exercise.[14] The same argument can be made for a put using its net exercise cash flow $F_t - K - C_t < 0$.[15] ■

Corollary. *From Theorem 4, $C_t = c_t$ and $P_t = p_t$, therefore all the results of this section are valid for American PFO.*

[14] Recall the convention of considering the payoff $F_\tau - K$ as cash settled, and the futures as initiated at the settlement price F_τ.

[15] In practice, the long can exercise an option during trading hours and also after market close until the time determined by the rules of the exchange. Moreover, he can also withdraw a submitted exercise order until that time. Therefore the *effective* decision time is after closing time. However, at this time, the resulting futures position cannot be closed. Therefore, the assumption of a perfectly liquid market is not satisfied. In the proof, I have assumed that the effective decision time is the *closing time* of the market. The proofs of Section 4 are robust to the above institutional detail.

Apropos Theorem 4, it is interesting to compare a PFO with a traditional-style option on futures. The latter can be optimally exercised before expiration when it is deep enough in the money. Early exercise's motivation lies in instantly collecting the intrinsic value in cash: the interests until expiration on the money received can preponderate against the loss of insurance. This is the case when the option is deep enough in the money because of both the larger amount of capital available for interest and the lower value of insurance. In the case of PFO, the intrinsic value (plus time value minus premium) has already been cashed in the form of variation margins. Therefore, there is no incentive to exercise early.

4 Non-optimality of early exercise in the presence of frictions

Up to this point, I have assumed that the market of options and futures is perfectly liquid, that there are no transaction costs and that the settlement prices are exactly equal to the true market price. In this section, I prove the suboptimality of early exercise of PFO without relying on those assumptions. As the framework becomes more general and realistic, the results will have useful implications for the design of options contracts by exchanges. In this section and the next, I deal exclusively with call options because the results and proofs for calls apply, mutatis mutandis, to puts.

Sufficient assumptions to derive the results of this section for PFO are:

Assumption 3 *Before expiration, there is always a positive probability that an option will be in the money at expiration, and a positive probability that, at expiration, it will be out of the money by more than the transaction costs TC_T of buying a futures contract at time T.*[16] *In the case of a call, this means that, for any time t such that*

[16] Without the Assumption 2, the suboptimality of early exercise cannot be proved. Nevertheless,

$t < T$, $\Pr_t(F_T > K) > 0$, *and* $\Pr_t(F_T < K - TC_T) > 0$.[17]

Assumption 4 *In determining the settlement price of a PFO and of the underlying futures contract during their lives, the exchange does not violate a minimal consistency condition between these two prices:*

$$C_t \geq (F_t - K)^+ . \tag{5}$$

In a market with frictions, this minimal consistency condition can be motivated by the following argument. If $C_t < (F_t - K)^+$, buying a futures contract would be strictly dominated by buying a call and exercising it on the same day, thus reaping the net exercise cash flow $(F_t - K)^+ - C_t$. (Variation margins starting next business day will be the same for either strategy.)

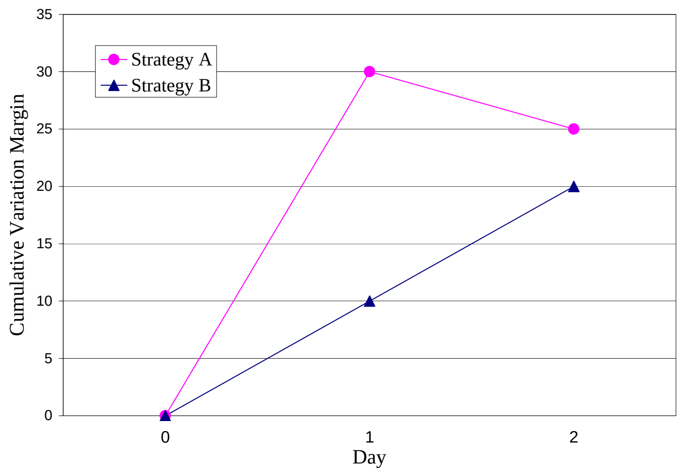
Transaction costs prevent the use of arbitrage arguments, which were the basis of the previous section. Instead, to prove the suboptimality of early exercise, I will rely on a novel dominance criterion, which is based on the gains processes of the PFO, of the underlying futures contract and of combinations of them. In our context of futures-style contracts, the gains process $\{G_t\}_t$ represents the evolution of the cumulative variation margins. I observe that G_t is *not* the value of a self-financing portfolio resulting from reinvesting the variation margins at an interest rate. It is the mere sum of the variation margins at and before t. For example, suppose that strategy C consists of buying and holding a pure futures call, then its gains process is:

$$G_t^C = (C_1 - C_0) + (C_2 - C_1) + ... + (C_t - C_{t-1}) = C_t - C_0.$$

even without this assumption, it can be proved that exercising only at expiration is an optimal strategy, in the sense that it is not dominated by any other. In other words, it cannot possibly hurt the buyer of a PFO to be bound to exercise only at expiration. This will become obvious in the proofs below.

[17] In the case of a put, Assumption 2 states that, for any time t such that $t < T$, $\Pr_t(F_T < K) > 0$, and $\Pr_t(F_T > K + TC_T) > 0$.

Figure 1: Dominance in Terms of Gains Processes



To help visualize the following Lemma and the supporting argument, Figure 1 displays the gains process of two strategies. The corresponding *single-day* variation margins are recorded in the following table.

	Strategy A	Strategy B	Differential Str.
Day 1	30	10	20
Day 2	-5	10	-15

Lemma 2 *Let G_t^A be the gains process of strategy A and let G_t^B be the gains process of strategy B. If G_t^A is not exceeded by G_t^B for any t, and there exists the possibility that G_T^A surpasses G_T^B, then strategy A dominates strategy B. More formally, for A's dominating B, it is sufficient that $\Pr\left(G_t^A \geq G_t^B\right) = 1$ for all $t < T$, and $\Pr\left(G_T^A > G_T^B\right) > 0$.*

This lemma is not entirely obvious because the condition $G_t^A \geq G_t^B$ allows for the possibility that, on particular days, the cash flow of strategy A be lower than that of strategy B (e.g., in day 2). To see that the lemma is always true in spite of the previous comment, note that, for each *negative* differential cash flow of strategy

A with respect to strategy B (e.g., -$15 in day 2), there will have been *previous positive* differential cash flows of equal or greater absolute value (e.g., $20 in day 1) so that $G_t^A \geq G_t^B$ for all t. As the positive differential cash flows lead, then, under non-negative interest rates, strategy A is still preferable. Negative interest rates are precluded by the absence of arbitrage.[18]

In a frictionless market, the exercise strategy is independent of the subsequent strategy of the investor; what matters are only the cash flows generated, and the value of the remaining position (always zero in that case), which can be unwound at will, and without incurring any transaction cost. This sat at the core of Theorem 4. However, in the presence of transaction costs, it is costly to unwind an unwanted position. As a result, the optimal strategy may vary per investor. I will analyze the exercise decision, first, assuming that the investor would like to hold the futures position in case of exercise (case A) and, second, assuming that he would prefer to close it (case B).

Case A

Assume that, in case of exercising a call, the investor does not intend to close the resulting futures position. He is willing to continue to bet on futures prices increasing. Here, no transaction costs are incurred in exercising the option or maintaining this position; this makes the proof for case A simpler than that for case B, and recommends the former as good practice for the latter. Even though transaction costs do not play a role for the particular traders considered in case A, they do play a role for others; therefore, the arguments of Section 3 no longer hold because the existence of transaction costs breaks their no-arbitrage foundation. In addition, illiquidity can hinder the determination of accurate settlement prices, making consistency condition

[18]Non-negative interest rates are also ensured in the absence of an interest rate market, which is equivalent to interest rates equaling zero. This comment is motivated by the claim made in the Introduction that the proofs of this article do not rely on the existence of an interest rate market.

(5) a necessary assumption to derive the results of the present section.

Theorem 5 *Assume that, if the investor exercised his pure futures call, he would choose to* keep *the resulting long futures position. Not exercising early is a dominating strategy.*

Proof. Let τ be the settlement time of a potential early-exercise date, and let G_t^E and G_t^N be the gains processes of the exercise and no-exercise strategies, respectively, defined for t such that $\tau \leq t \leq T$, and including the cash flows of the decision day onward. The gains process of the exercise strategy is

$$\begin{aligned} G_t^E &= (C_\tau - C_{\tau-1}) + (F_\tau - K - C_\tau) + (F_t - F_\tau) \\ &= F_t - K - C_{\tau-1}, \end{aligned} \tag{6}$$

where $C_\tau - C_{\tau-1}$ is the variation margin of the call on the exercise day, $F_\tau - K - C_\tau$ is the net exercise cash flow, and $F_t - F_\tau$ is the cumulative margin of the futures resulting from exercise. The gains process of the no-exercise strategy, including the variations margin of the decision day onward, is

$$G_t^N = C_t - C_{\tau-1} \tag{7}$$

$$\begin{aligned} &\geq (F_t - K)^+ - C_{\tau-1} \\ &\geq F_t - K - C_{\tau-1} + (K - F_t)^+ , \end{aligned} \tag{8}$$

where the inequality arises from the exchange's respecting consistency condition (5).

Comparing (6) with (8), we can see that $G_t^N \geq G_t^E$ for all $t \geq \tau$ with probability one, and $G_T^N > G_T^E$ with positive probability by Assumption 3. Therefore, the no-exercise strategy dominates the exercise strategy by Lemma 2. ■

In the case of a put, it can be shown, using similar arguments, that exercising with the intention of keeping the resulting short futures, is dominated by merely keeping the option alive.

Case B

For this case, we can imagine an investor who is satisfied with the intrinsic value reached by a call and who, fearing its erosion by a drop of the price of the underlying, is tempted to exercise and close the resulting futures.

Theorem 6 *Assume that, if the investor exercised his pure futures call, he would choose to* close *the resulting long futures position. Then, shorting futures and not exercising is a dominating strategy.*

Proof. Let τ be, as above, the settlement time of a potential early-exercise date and τ^- a time during trading hours of that day. Should the investor exercise the call and sell futures at time τ^- during the trading hours (in order to lock in the intrinsic value at that time), the gains process is

$$\begin{aligned} G_t^E &= (C_\tau - C_{\tau-1}) + (F_\tau - K - C_\tau) - (F_\tau - F_{\tau^-}) - TC_\tau \\ &= F_{\tau^-} - K - C_{\tau-1} - TC_\tau, \end{aligned} \tag{9}$$

where $-(F_\tau - F_{\tau^-})$ is the variation margin at τ of the short futures opened at τ^-, and TC_τ are the transaction costs of selling futures. I include neither the sum $-(F_t - F_\tau)$ of the variation margins of this short futures position nor the sum $(F_t - F_\tau)$ of the variation margins of the long futures generated by exercise because the exchange closes the two positions after computing the variation margins at τ. Note that t is not present in (9), which means that the cumulative cash flow is constant for $t \geq \tau$. This needs be the case because all positions are closed at τ.

If the investor only sells a futures contract at the same time τ^- and does not exercise the option, the gains process G_t^N for $\tau \leq t < T$ is

$$
\begin{aligned}
G_t^N &= (C_t - C_{\tau-1}) - (F_t - F_{\tau^-}) - TC_\tau \qquad (10) \\
&\geq \big((F_t - K)^+ - C_{\tau-1}\big) - ((F_t - K) - (F_{\tau^-} - K)) - TC_\tau \\
&\geq F_{\tau^-} - K - C_{\tau-1} - TC_\tau + (K - F_t)^+ , \qquad (11)
\end{aligned}
$$

where the weak inequality arises from the exchange respecting consistency condition (5) and where $(K - F_t)^+$ in (11) is the result of $(F_t - K)^+ - (F_t - K)$ from the previous line.

Comparing (9) with (11), we realize that $G_t^N \geq G_t^E$ for $\tau \leq t < T$ with probability one. No assertion is made for $t = T$ because other transaction costs need to be considered.

Recall that the no-exercise strategy kept the long call, and added a short futures contract in order to lock in the gains obtained at τ^-. If the call expires out of the money by more than the transaction costs TC_T of buying a futures contract, $F_T < K - TC_T$, the short futures contract of the existing strategy is cancelled by buying futures; this locks in extra profits and eliminates any exposure to future losses. If $F_T \geq K - TC_T$, then exercising the call generates long futures that cancel the short futures of the existing position and the buyer derives no additional gains.

In the case where $F_T < K - TC_T$, transaction costs TC_T are incurred to close the short futures position at T:

$$
\begin{aligned}
G_T^N &= (C_T - C_{\tau-1}) - (F_T - F_{\tau^-}) - TC_\tau - TC_T \\
&= (0 - C_{\tau-1}) - (F_T - F_{\tau^-}) - TC_\tau - TC_T \\
&= F_{\tau^-} - K - C_{\tau-1} - TC_\tau + (K - TC_T - F_T) . \qquad (12)
\end{aligned}
$$

The expression (12) of G_T^N is greater than the expression (9) of G_T^N when $F_T < K - TC_T$, which happens with positive probability by Assumption 3. In the case where $F_T \geq K - TC_T$, it is optimal to exercise the option at expiration. In this case,

$$\begin{aligned} G_T^N &= (C_T - C_{\tau-1}) - (F_T - F_{\tau^-}) - TC_\tau \\ &= ((F_T - K) - C_{\tau-1}) - (F_T - F_{\tau^-}) - TC_\tau \\ &= F_{\tau^-} - K - C_{\tau-1} - TC_\tau \end{aligned}$$

and $G_T^N = G_T^E$.

As we have seen, $G_t^N \geq G_t^E$ for $\tau \leq t < T$, $G_T^N = G_T^E$ when $F_T \geq K - TC_T$, and $G_T^N > G_T^E$ for $F_T < K - TC_T$, which happens with a positive probability. Therefore, the no-early-exercise strategy dominates the early-exercise strategy by Lemma 2. ■

Another dominant strategy to the early-exercise strategy can be to short a call with the same strike and expiration. However, the liquidity of a call option with the same strike and expiration cannot always be guaranteed and is most probably lower than that of the futures. Liquidity of the futures is not guaranteed either; this could be thought to compromise the plausibility of the dominating strategy of the previous theorem. Nevertheless, if the strategy of exercising the option and selling futures is feasible, then so is the dominating strategy of just selling futures.

In the case of a put, a buyer tempted to exercise early and buy futures to close his position can be shown to be better off by merely buying futures, and not exercising. The proof is similar to that for a call.

In order to highlight the importance of the result of this section, I note that early exercise may be optimal for traditional-style options on a non-dividend-paying stock, especially when they are deep in the money, and time to expiration is short. This remark contradicts what derivatives textbooks teach, which is true only in frictionless markets. In the real world, the bid premium can be below the intrinsic value; therefore, a trader who expects the stock's price to decrease might be better off by exercising the option early rather than selling it. It may be contended that, instead of exercising the option, the trader can short sell the stock. However, this is risky because you can short the stock only on an upstick or a zero-plus tick, and the price may drop substantially before the trade can be executed. (An uptick is a transaction occurring at a price above that of the previous transaction. A zero-plus tick is a transaction at the same price as that of the preceding trade, but higher than that of the last trade at a different price.) In addition, when you short sell stock, you post collateral of at least 150% of the value of the short sold stock. Then, the collateral must not fall below a maintenance margin of 130%, or 100% plus $5 times the number of shares, whichever the greater.[19] Once the stock has been short sold, the broker has the discretion to demand that the position be closed out at any time. This may happen when there are many investors selling or short selling the stock, and the broker runs out of inventory and cannot procure it from other brokers. Such a scenario might even prevent short selling in the first place.

5 Performance bond

In this section, I analyze the implications of collateral requirements on the early exercise strategy.

Apart from variation margins, exchanges require from traders an additional margin, called a "performance bond", under which they have to deposit sufficient collateral to cover the one-day value at risk (VAR) of their positions. VAR is computed at a portfolio level, where a portfolio groups positions in a futures contract with all options on that contract. The calculations consider a number of scenarios with dif-

[19] To be able to post collateral, you need to have already opened a margin account with your broker.

ferent combinations of the next day's futures price and implied volatility. The VAR of the portfolio is the loss under the worst case scenario.[20]

For a time t, the next day's value of a call in the worst case scenario of the portfolio is

$$\hat{C}_{t+1} \equiv C\left(\hat{F}_{t+1},\ \hat{\sigma}_{t+1},\ t+1, ...\right),$$

where $\hat{F}_{t+1}$ and $\hat{\sigma}_{t+1}$ are the futures price and the volatility in the worst case scenario of the portfolio, respectively, and where $C(\ )$ is the pricing function used by the exchange.

The *contribution* of a long call to the *VAR of a portfolio* is

$$VAR_t^C = -\left(\hat{C}_{t+1} - C_t\right). \tag{13}$$

In the proofs of the Appendix, I will also make use of the contribution of a long futures contract,

$$VAR_t^F = -\left(\hat{F}_{t+1} - F_t\right), \tag{14}$$

and of a short futures contract,

$$VAR_t^{-F} = -\left(F_t - \hat{F}_{t+1}\right). \tag{15}$$

Each of these contributions can be positive or negative; however, the VAR of the portfolio in the worst case scenario is always positive. The VAR of the portfolio is equal to the sum of the contributions of the individual positions. A negative contribution means that the corresponding position acts as a hedge thereby reducing the VAR of the portfolio.

[20] See Eurex (2003) for further details. Most of the medium to large exchanges in the world use SPAN (Standard Portfolio Analysis of Risk), a system designed by the CME, which computes performance bond requirements using the method described. This system can handle both traditional-style and futures-style options.

The required value of performance bond at each time t equals the VAR of the corresponding portfolio. The performance bond and the variation margin described in Section 2 differ not only in the way they are computed. While the variation margin paid by one trader is passed to another, the performance bond posted by a trader is kept by the clearing house. Unlike the variation margin, which must be deposited in cash, the performance bond may also take the form of liquid securities, or bank guarantees. The clearing house determines the list of the eligible securities, the haircut to their market value, and the list of eligible banks. While the required value of the performance bond is a state variable (a stock), a variation margin is a flow. However, the performance bond can be subtracted from the cumulative variation margin because the latter is also a state variable. The result of this subtraction is the net liquidity accumulated by a trader, should the performance bond be posted in the form of liquid securities, and should haircuts be ignored.

The effect of the performance bond on the exercise strategy cannot be analyzed separately from the effect of the variation margins. For example, if an investor exercises a call, this position is replaced by a futures contract. Although VAR_t^F will most probably be greater than VAR_t^C, we cannot make a general statement. A scenario assuming a sizeable drop in the implied volatility could result in the opposite inequality. Therefore, exercising a call could bring about a decrease of collateral. However, it will be demonstrated that this potential decrease of collateral will always be compensated by a negative net exercise cash flow. Moreover, variation margins or net exercise cash flows are more important than changes in required performance bond because the latter can be posted not only in cash, but also in traded securities. Therefore, it is irrelevant to consider the changes in collateral in isolation. To consider the performance bond from this perspective, I will focus on a "modified gains process" to be denoted by H_t, and defined as the gains process G_t

of the previous section minus the requirements for collateral:

$$H_t \equiv G_t - VAR_t. \tag{16}$$

Considering the cumulative variation margin and the required collateral, both indexed by t, will ensure that conclusions account for variation margins and collateral requirements of future days, and not only of the exercise day.

The previous section established that, before expiration, the no-exercise strategy (with gains process G_t^N) always dominates the exercise strategy (with gains process G_t^E). Theorem 7 will show that considering H_t^N and H_t^E does not dispel that dominance.

Lemma 3 *Let* $\hat{G}_{t+1} \equiv \hat{G}\left(\hat{F}_{t+1},\ \hat{\sigma}_{t+1},\ t+1, ...\right)$ *be the gains process at* $t+1$ *but computed at* t*, using the futures price* $\hat{F}_{t+1}$ *and the implied volatility* $\hat{\sigma}_{t+1}$ *of the worst case scenario considered by the performance bond system. Then,*

$$H_t = \hat{G}_{t+1}.$$

Proof. VAR_t is the loss at $t+1$ in the worst case scenario; therefore, subtracted from the gains process at t, yields the gain process at $t+1$ in the worst case scenario, $\hat{G}_{t+1} = G_t - VAR_t$, which equals H_t by definition (16). In the Appendix, I offer separate proofs of Lemma 3 for both the holding and the exercise strategies in each of the cases A and B of the previous section. ■

Theorem 7 *The introduction of the performance bond system does not alter the dominance of the no-exercise strategy over the exercise strategy before expiration.*

Proof. Recall that, in any of the two cases of the last section, G_t^E was not greater than G_t^N in any possible scenario. $\hat{G}_{t+1}^E$ and $\hat{G}_{t+1}^N$ are gain processes under a restricted set of those scenarios. Therefore, $\hat{G}^E_{t+1}$ never exceeds $\hat{G}^N_{t+1}$ and, by Lemma 3, H^E_t is never greater than H^N_t. This is enough to maintain the dominance of the no-exercise strategy that was obtained in the previous section in terms of G^E_t and G^N_t, which are *cash* accumulated gains. (In a situation like $H^N_t = H^E_t$ and $G^N_t > G^E_t$, the no exercise strategy dominates the exercise strategy because the investor prefers cash to collateral.) The Appendix offers separate proofs of Theorem 7 for each of the cases A and B of the previous section. ■

6 Discussion on contract design

Because it is never optimal to exercise a PFO before expiration, there is no reason to preserve the American feature in the specification of this contract. This conclusion is robust to transaction costs, illiquidity, and performance bond requirements. It does assume that investors are able to analyze the alternative strategies presented above, that they actually do so, and that they prefer more to less. Failing any of these assumptions, exchanges might include the American feature out of marketing considerations. However, Easton (1997) states that early exercise in the SFE is very rare.[21] It is possible that some investors assign a subjective value to the American feature at the moment of a trade but, on intending to exercise, are advised by their brokers to reconsider. This would reconcile the exchanges' maintaining the American feature with an observed low number of exercises. Finally, a reason for maintaining the American feature is the low cost of keeping with tradition.[22]

[21] He reports number of exercises in 1994 for the four major option contracts at that time. I collected the corresponding volumes for the same year. The number of exercises and the corresponding volumes (in parenthesis) are the following: 6 (833,667) for the All-Ordinaries Share Price Index Contract, 40 (943,749) for the 90-Day Bank Accepted Bill Contract, 40 (507,252) for the 3-Year T-Bond Contract and 300 (800,263) for the 3-Year (10-Year) T-Bond Contract. I was not able to obtain further information from this or other exchanges.

[22] The cost of keeping the American feature derives from the exercise procedure, which includes assigning the resulting futures contract to an option writer. In fact, the exchange could even avoid the latter by assigning the option to itself. This would not impose any risk on the exchange; rather,

If an option contract is illiquid, the exchange can still use futures-style margining by equating settlement prices to the outputs of a pricing model. Moreover, if the exchange does not dare to make the variation margins, which are settled in cash, depend on a model, whose calibration might be subject to dispute, the intrinsic value can be utilized as the settlement price. In this case, much of the benefit of futures-style margining would be retained when an option is in the money. Interestingly, even if exchanges are not precise in the determination of the settlement premia, the suboptimality of early-exercise still holds, provided they comply with consistency condition (5), which is satisfied by the intrinsic value and the output of any reasonable pricing model.

7 Conclusion

Without making any assumption about futures price or interest rate dynamics, this paper has shown that, in a no-arbitrage setting, the premium of a PFO before expiration always exceeds its intrinsic value, and that it is never optimal to exercise before expiration. The paper has also proved put-call parity, and equality of time values of puts and calls with the same strike and expiration.

A novel dominance criterion corroborates that early exercise is never optimal, even in the presence of transaction costs, illiquidity, or performance bond requirements, as long as the exchange determines option settlement prices satisfying a minimal rationality constraint. Consequently, in a market of sufficiently educated participants, including the American feature among the specifications of a PFO contract is of no avail. Finally, it is important to contrast the full generality and realism

it provide it with a free option. For example, if a trader exercises a call early, he gets a long futures position. The exchange, acting as his counterparty assumes a *short futures position*. As the short call position remains open, the exchange keeps a *long call option* against the writer. Thus, it effectively acquires, at zero premium, a portfolio equivalent to a long put with the same strike of the call. A similar argument holds if a put is exercised early.

of the suboptimality of early exercise of futures-style options to the limitation to a frictionless market of the parallel result for calls on a non-dividend paying stock.

Appendix

Proof of Theorem 1

In order to make a recursive argument, assume that, at some time $t + 1$,

$$c_{t+1} - p_{t+1} = F_{t+1} - K. \tag{17}$$

Now I will show that it follows that, at time t,

$$c_t - p_t = F_t - K.$$

If

$$c_t - p_t < F_t - K, \tag{18}$$

then the following strategy is an arbitrage: sell futures, buy a call and sell a put at closing time t. No cash flow is generated that day because, in a frictionless market, the settlement price coincides with the closing price. The variation margin at $t + 1$ is

$$\begin{aligned} & (F_t - F_{t+1}) + (c_{t+1} - c_t) + (p_t - p_{t+1}) \\ = \; & F_t - c_t + p_t + \underbrace{c_{t+1} - p_{t+1} - F_{t+1}}_{=-K \text{ by } (17)} \\ = \; & (F_t - K) - (c_t - p_t) > 0, \end{aligned}$$

where the inequality follows from (18). This is an arbitrage because there is zero investment at time t and a positive certain cash flow at time $t + 1$. Therefore, (18) cannot be true. If $c_t - p_t > F_t - K$, a symmetric arbitrage can be made, so this inequality cannot be true either.

Therefore, I have shown that

$$c_{t+1} - p_{t+1} = F_{t+1} - K \implies c_t - p_t = F_t - K.$$

To complete the argument, it is enough to note that put-call parity is trivially satisfied at expiration time T:

$$c_T - p_T = (F_T - K)^+ - (K - F_T)^+ = F_T - K.$$

Proof of Lemma 1

$C_T = (F_T - K)^+ \geq 0$ and, from Assumption 2, $\Pr_t \left(C_T = (F_T - K)^+ > 0 \right) > 0$, therefore $\Pr_t (C_{T-1} > 0) = 1$. (If $C_{T-1} \leq 0$, buying a call at $T-1$, the closing time of the day previous expiration, is an arbitrage. The cash flow at $T-1$ is 0, and the variation margin at T, $C_T - C_{T-1}$, can be positive and is never negative, which implies an arbitrage.) To go on backward, the following recursive argument can be followed: If $\Pr_t (C_{t+1} > 0) = 1$, then $C_t > 0$. (If not, buying a call is an arbitrage.) For a put, the structure of the argument is identical.

Proof of Lemma 3

The following case A and case B correspond to the cases of Section 4. I prove $H_t = \hat{G}_{t+1}$ for the exercise and the holding strategy of each of the cases.

Case A

If the investor decides to exercise the option and keep the resulting long futures position, H_t^E equals G_t^E, computed in (6) and repeated here in (19), minus the performance bond (14) for the futures position:

$$
\begin{aligned}
H_t^E &= G_t^E - VAR_t^E = G_t^E - VAR_t^F \\
&= (F_t - K - C_{\tau-1}) \qquad (19) \\
&\quad + \left(\hat{F}_{t+1} - F_t\right) \\
&= \hat{F}_{t+1} - K - C_{\tau-1} \qquad (20) \\
&= \hat{G}_{t+1}^E.
\end{aligned}
$$

Since G_t^E equals (19), it is clear that (20) is $\hat{G}_{t+1}^E$.

If the investor decides not to exercise, H_t^N equals G_t^N, computed in (7) and repeated here in (21), minus the performance bond (13) for the option:

$$
\begin{aligned}
H_t^N &= G_t^N - VAR_t^N = G_t^N - VAR_t^C \\
&= (C_t - C_{\tau-1}) \qquad (21) \\
&\quad + \left(\hat{C}_{t+1} - C_t\right) \\
&= \hat{C}_{t+1} - C_{\tau-1} \qquad (22) \\
&= \hat{G}_{t+1}^N.
\end{aligned}
$$

While G_t^N equals (21), it is clear that (22) is $\hat{G}_{t+1}^N$.

Case B

If the investor decides to exercise early and close the resulting futures position, there is no collateral to subtract; therefore, H_t^E equals G_t^E, computed in (9) and repeated here in (23):

$$\begin{aligned} H_t^E &= G_t^E - VAR_t^E = G_t^E - 0 \\ &= F_{\tau^-} - K - C_{\tau-1} - TC \qquad (23) \\ &= \hat{G}_{t+1}^E. \end{aligned}$$

Apart from being G_t^E, (23) is also G_{t+1}^E under any scenario ($\hat{G}_{t+1}^E$ in the worst case scenario) because the position was closed at τ, which is reflected in the fact that (23) is not indexed by t.

If the investor keeps the option position open and sells futures to lock in the intrinsic value, H_t^N equals G_t^N, computed in (10) and repeated here in (24), minus the performance bond requirement of the combined position, whose components are (13) and (15):

$$\begin{aligned} H_t^N &= G_t^N - VAR_t^N = G_t^N - VAR_t^C - VAR_t^{-F} \\ &= (C_t - C_{\tau-1}) - (F_t - F_{\tau^-}) - TC \qquad (24) \\ &\quad + \left(\hat{C}_{t+1} - C_t\right) - \left(\hat{F}_{t+1} - F_t\right) \\ &= \left(\hat{C}_{t+1} - C_{\tau-1}\right) - \left(\hat{F}_{t+1} - F_{\tau^-}\right) - TC \qquad (25) \\ &= \hat{G}_{t+1}^N. \end{aligned}$$

Since G_t^N equals (24), it is clear that (25) is $\hat{G}_{t+1}^N$.

Proof of Theorem 7

Case A

From (22),

$$\begin{aligned} H_t^N &= \hat{C}_{t+1} - C_{\tau-1} \\ &\geq \left(\hat{F}_{t+1} - K\right)^+ - C_{\tau-1} \\ &\geq \hat{F}_{t+1} - K - C_{\tau-1} + \left(K - \hat{F}_{t+1}\right)^+ . \end{aligned} \tag{26}$$

Comparing (26) and (20), it is clear that H_t^E is never greater than H_t^N.

Case B

From (25),

$$\begin{aligned} H_t^N &= \left(\hat{C}_{t+1} - C_{\tau-1}\right) - \left(\hat{F}_{t+1} - F_{\tau^-}\right) - TC \\ &\geq \left(\left(\hat{F}_{t+1} - K\right)^+ - C_{\tau-1}\right) - \left(\left(\hat{F}_{t+1} - K\right) - (F_{\tau^-} - K)\right) - TC \\ &\geq F_{\tau^-} - K - C_{\tau-1} - TC + \left(K - \hat{F}_{t+1}\right)^+ . \end{aligned} \tag{27}$$

Comparing (27) and (23), we conclude that H_t^E never exceeds H_t^N.

References

Barone-Adesi, G. & Whaley, R. (1987), 'Efficient analytic approximation of American option values', *Journal of Finance* **42**, 301–320.

Black, F. (1976), 'The pricing of commodity contracts', *Journal of Financial Economics* **3**, 167–179.

Chen, R.-R. & Scott, L. (1993), 'Pricing interest rate futures options with futures-style margining', *Journal of Futures Markets* **13**, 15–22.

Cox, J. C., Ingersoll, J. E. & Ross, S. A. (1981), 'The relationship between forward prices and futures prices', *Journal of Financial Economics* **9**, 321–346.

Cox, J. C., Ingersoll, J. E. & Ross, S. A. (1985a), 'An intertemporal general equilibrium model of asset prices', *Econometrica* **53**, 363–384.

Duffie, D. (1989), *Futures Markets*, Prentice-Hall.

Duffie, D. & Stanton, R. (1992), 'Pricing continuously resettled contingent claims', *Journal of Economic Dynamics and Control* **16**, 561–573.

Easton, S. A. (1997), 'Put-call parity with futures-style margining', *Journal of Futures Markets* **17**, 215–227.

Eurex (2003), *Risk Based Margining*, Eurex Publications.

Eurex (2004), *Products 2004*, Eurex Communications.

Frey, R. & Sommer, D. (1998), 'A generalization of the Geske-formula for compound options to stochastic interest rates is not trivial - A note', *Journal of Applied Probability* **35**, 501–509. AMS.

Kuo, C. K. (1993), 'The valuation of futures-style options', *The Review of Futures Markets* **10**, 481–487.

Kutner, G. W., Porter, D. C. & Thatcher, J. G. (2001), 'The proposed introduction of futures-style margining in the U.S.: An Australian comparison', *Journal of Financial Research* **24**, 239–259.

Lieu, D. (1990), 'Option pricing with futures-style margining', *Journal of Futures Markets* **10**, 327–338.

Merton, R. C. (1973), 'Theory of rational option pricing', *Bell Journal of Economics and Management Science* **4**, 141–183.

Satchell, S., Stapleton, R. C. & Subrahmanyam, M. G. (1997), 'The pricing of marked-to-market contingent claims in a no-arbitrage economy', *Australian Journal of Management* **22**.

Twite, G. (1996), 'The pricing of SPI futures options with daily futures style margin payments', *Australian Journal of Management* **21**, 139–157.

White, A. J. (1999), *Pricing Options with Futures-Style Margining: A Genetic Adaptive Neural Network Approach*, Garland Publishing, N.Y.

Improving the Design of Treasury-Bond Futures Contracts

Rodolfo A. Oviedo*

First Version: November 2001
This Version: September 11, 2003

Abstract

In bond futures contracts, the seller can choose which bond to deliver from a basket of eligible issues. A Conversion Factor System is used by the CBOT and other exchanges in the world in order to adjust the futures invoice price according to the bond chosen for delivery. This system was designed with the objective of making the futures invoice prices for the different eligible bonds close to their corresponding spot market prices. However, the poor performance of this system in achieving this objective is very well known. In this paper, I propose an alternative method for computing futures invoice prices: the True Notional Bond System. I show that this method meets the aforesaid objective much better than the Conversion Factor System. In addition, it makes futures contracts more effective and easier to use as a tool for risk management.

*I am grateful to the Institut de Finance Mathématique de Montréal (IFM^2) for generous financial support. I would like to thank Peter Christoffersen, Jan Ericsson and Kris Jacobs for helpful comments. Any remaining inadequacies are mine alone. Correspondence to: Rodolfo Oviedo, Faculty of Management, McGill University, 1001 Sherbrooke Street West, Montreal, Canada H3A 1G5; Tel: (514) 398-4000 ext. 00854#; Fax: (514) 398-3876; E-mail: rodolfo.oviedo@mail.mcgill.ca.

Contents

1 Introduction

Specifications of most government bond futures contracts allow the seller to choose at expiration which issue to deliver among a list of eligible bonds. Typically, this list includes bonds within a maturity range issued by the government of a single country. Therefore, when two traders enter into a futures contract, the bond to be delivered is not yet known; they only know that it will be one from the list corresponding to that contract; only *at expiration* will the seller choose one bond from this list. The main objective of the multi-asset feature is to facilitate the delivery of the underlying, thus avoiding liquidity problems in the spot market that could make it difficult or unduly expensive for the seller to make delivery. The multi-asset design of Treasury-bond futures contracts described here has been adopted by, among others, the CBOT, EUREX, LIFFE, MATIF, MEFF, and the Montreal Exchange. In this paper, I focus on the Treasury-bond futures contract traded on the CBOT, one of the most successful derivative contracts in the world.

For simplicity, I will assume that the following events occur simultaneously at time T: (i) the futures contract stops trading, (ii) the last settlement price becomes known, (iii) the short chooses the bond to deliver, (iv) the seller delivers, and (v) the buyer pays for the bond. In futures contracts on a single underlying, the short party receives the last settlement price F_T as a payment for the commodity he delivers; in other words, the *futures invoice price* upon delivery is F_T. For multi-asset T-Bond futures, however, eligible bonds differ in maturity and coupon and, therefore, they have different spot market values. Using the raw F_T as the futures invoice price would make it far too disadvantageous for the short trader to deliver a bond other than the cheapest, thus rendering the ability to choose the bond to deliver useless. To mitigate this problem, the futures invoice price FIP^X is made a function not only of F_T but also of the particular bond X chosen by the seller to deliver:

$$FIP^X = g\left(F_T, X\right) \tag{1}$$

The objective of this *futures invoice price function* is to make the futures invoice prices of the different bonds as close as possible to their corresponding spot market prices. When this objective is met, the demand of bonds for the purpose of delivery is less concentrated on a single issue, the overpricing of this bond and of the futures contract is minimized, and deliberately provoked shortages of that issue are discouraged.

The ability of the short party to choose which bond to deliver is referred to in the literature as the *quality option*. Obviously, the value of the quality option crucially depends on the form of the futures invoice price function. A substantial amount of research has been done in the valuation of the quality option embedded in Treasury Bond Futures Contracts.[1] However, to the best

[1] A sample of the papers studying the valuation of the quality option include Cox, Ingersoll and Ross (1981), Gay and Manaster (1984), Kane and Marcus (1986), Kamara and Siegel (1987), Livingston (1987), Hegde (1988), Boyle (1989), Hegde (1990), Hemler (1990), Carr

of my knowledge, no paper has addressed an even more fundamental research question: Is the current functional form $g(\cdot,\cdot)$ optimal according to the aforesaid objective? The literature has noticed the drawbacks of the current functional form, but no attempt has been made to provide an alternative[2]. The purpose of this paper is to propose a novel functional form $g(\cdot,\cdot)$ for the futures invoice price function (1), that brings the futures invoice prices of the different eligible bonds dramatically closer to their corresponding spot market prices.

Currently, the futures invoice price is determined according to the Conversion Factor System (CFS), whose futures invoice price function is $g(F_T, X) \equiv F_T\, CF^X$,[3] where CF^X is a *predetermined* constant referred to as the conversion factor, which is different for each bond X. Ideally, when time T arrives, the futures invoice price $F_T\, CF^X$ for each bond X would equal its spot market price S_T^X. Then, the short would be indifferent between delivering any of the eligible bonds. But this is true only for very particular values of the conversion factors; and these ideal values depend on market conditions at T. However, in the CFS, the conversion factors are precomputed quantities[4], seldom close to the ideal ones. The alternative proposed in this paper can be *thought of* as allowing conversion factors to be computed as a function of market conditions at T. This allows the futures invoice price to be markedly closer to the spot market price of each eligible bond and, as a result, the objectives of the futures invoice price function are much better met.

In the new method I propose in this paper, the futures invoice price for any eligible bond X is computed by discounting the remaining cash flows of X at the "yield implied by F_T". The latter is the yield y^{NB} of a *notional* bond "issued" at T, whose tenor and coupon are fixed in the specifications of the futures contract, and whose price is defined to be F_T. I refer to this way of computing the futures invoice price as the True Notional Bond System (TNBS). According to the previous explanation, the futures invoice price is a function of F_T and X, the same inputs used by the CFS. Therefore, the difference between the two systems is just a different form $g(\cdot,\cdot)$ of the futures invoice price function (1).

Key in the TNBS is the yield y^{NB} of the notional bond, which, under no-arbitrage, will be shown to equal the highest among the yields of all eligible issues. Therefore, when the yields of the eligible bonds are similar, discounting all bonds at y^{NB} produces futures invoice prices that are close to the corresponding spot market prices. Given that all eligible bonds must belong to a particular range of maturities, it is reasonable to assume that their yields are similar. In the

and Chen (1997), Yu (1999), Magdon-Ismail, Atiya and Abu-Mostafa (2000) and Vidal Nunes and Ferreira de Oliveira (2003). Some of these papers also deal with hedging in the presence of the quality option.

[2] See, for example, Kane and Marcus (1986), Johnston and McConnell (1989), or Schulte and Violi (2001). Kane and Marcus (1984) propose a minor variation to the CF System. However, their goal is to increase the hedging effectiveness of the futures contract, which is different from the main objective of this paper.

[3] The actual formula is $g(F_T, X) \equiv F_T\, CF^X + AI_T^X$, where AI_T^X are the accrued interests of the bond X at time T.

[4] The conversion factor of each bond is fixed using a formula that depends only on the cash flows maturing after T, with no input related to market conditions at T.

ideal case of a flat yield curve, *whatever its level be*, futures invoice prices would perfectly coincide with spot market prices. For the CFS, a flat yield curve is not enough; *6% is the only level*[5] that obtains the desired equality of futures invoice prices and spot market prices. Therefore, even when, in reality, yields curves are not flat, it is not surprising that the empirical part of this paper reports a clear superiority of the TNBS over the CFS. The similarity among the yields of the eligible bonds is responsible for the good empirical performance of the TNBS. The situation is somewhat analogous to that of duration-based strategies that rely on the assumption that the yield curve is flat: these strategies perform well empirically even when the yield curves are not flat. A refinement of the TNBS is suggested at the end of the paper to further improve its behavior under non-flat yield curves.

The paper is organized as follows: Section 2 explains the criterion to compare the TNBS to the CFS, Section 3 states no-arbitrage conditions common to both, Section 4 reviews the CFS, Section 5 develops the TNBS, Section 6 empirically compares the two systems, Section 7 outlines some possible extensions, and Section 8 concludes.

2 Criterion for Evaluating Alternative Futures Invoice Price Functions

The optimal bond for the seller to deliver is referred to as the *cheapest-to-deliver* (*CTD*); any other eligible bond will be referred to as an *alternative bond*. The *preferred* form $g(\cdot, \cdot)$ of the futures invoice price function (1) will be the one with *lower losses* for delivering alternative bonds. The next section will show that this criterion is equivalent to the one mentioned before: The preferred form of (1) is the one that makes the futures invoice prices of the different bonds closer to their corresponding spot market prices at T. Next, I discuss four advantages that result when the short party faces small losses in delivering alternative bonds.

First, when those losses are small, the short has an economically real possibility of delivering different bonds.[6]

[5] It used to be 8%, the reference rate for contracts that expired before March, 2000.

[6] Suppose that the *marginal traders*, who determine the last settlement price F_T (and consequently the futures invoice price of all eligible bonds by (1)), *do not own* any of the eligible bonds and that, therefore, they need to buy them at the *ask spot price*. For the present argument, I will denote by *CTD* the bond of optimal delivery *by the aforementioned marginal traders who do not own eligible bonds*; the rest will be the referred to as the alternative bonds. Although those traders will make a loss if they deliver a bond other than the *CTD*, this might *not* be the case for traders who already own one of the alternative bonds. Consider, for example, a trader who is short in the future and hedged with a long spot position in one of the eligible bonds. He may want to liquidate his position by delivering this bond. The reason is that his relevant cost of delivery is the bid spot price (lower that the ask) and, therefore, he might find cheaper to deliver the bond he owns. However, if the loss from delivering an alternative bond is big, then considering bid or ask spot prices will make no difference and, therefore all the demand will be concentrated in the *CTD*.

Second, when the loss of delivering alternative bonds is small, the price pressure on the spot price of the *CTD* (due to the short traders' trying to buy it) is small. The reason is that, when the loss of delivering an alternative bond is small, a minor increase in the price of the *CTD* makes this loss disappear. This, in turn, allows the short to optimally deliver another bond and, as a result, the price pressure disappears. Therefore, the upper bound of the possible overpricing of the *CTD* is small.[7]

Third, the resulting small upper bound of the overpricing of the *CTD* prevents a material overpricing of the futures contract.[8]

Fourth, the long traders have little incentive to try to squeeze the market of the *CTD*, i.e. to deliberately provoke a shortage of the *CTD*. Provoked shortages of the *CTD* may increase its price and, consequently, increase the futures price. The intention is to generate an extraordinary profit in the futures market for the long at the expense of the short, or to sell the bonds at a premium taking advantage of the demand by the short futures traders. As explained in the two previous paragraphs, the potential overpricing of the *CTD* and the future are small when so are the losses of delivering alternative bonds.[9]

Because of these reasons, a relevant metric for comparing the performance of alternative functional forms $g(\cdot,\cdot)$ for (1) is the loss generated by the delivery of alternative bonds.[10]

3 Common No-Arbitrage Conditions

This section provides no-arbitrage conditions applicable both to the Conversion Factor System and to the True Notional Bond System and, in general, to any functional form $g(\cdot,\cdot)$ in (1). These conditions will be used later to derive the last settlement price F_T and to identify the *CTD* in each of the aforementioned systems. In addition, the equivalence of the two criteria mentioned in the first paragraph of the previous section is shown.

[7] As Jordan and Kuipers (1997) point out, "the impact of a derivative security *on the price of its underlying asset* is a central research and public policy issue." They show how the futures contract can significantly distort the spot price of the *CTD*. Also, they remark "the importance of contract design" and specifically mention the conversion factor system as a cause of the aforementioned distortions. However, they do not propose any alternative to the conversion factor system.

[8] The connection between the futures price and the spot price of the *CTD* will be shown later: the former is an increasing function of the latter, both in the CF System and in the True NB System.

[9] Schulte and Violi (2001) consider the effect of the futures contracts on the European spot market. They state that "The tremendous level of activity in the EUREX contract has raised concerns about the risk of a shortage in the cheapest to deliver". They argue that, in order to avoid squeezes, it is important that it not be easily forecastable which bond will become the *CTD* at expiration. Merrick, Naik and Yadav (2002) investigate a well-publicized market manipulation episode: an attempted delivery squeeze in a bond futures contract traded in London.

[10] Possible alternative metrics are the upper bound of the overpricing of the *CTD* or of the future. Computations using these metrics are not reported but yield similar results both qualitatively and quantitatively.

The profit resulting from the short's delivering a particular bond X is:

$$Profit(X) = FIP^X - S_T^X \tag{2}$$

where FIP^X is the futures invoice price to be paid by the long to the short, and S_T^X is the full spot market price at T of bond X.

The optimal bond for the seller to deliver, referred to as the cheapest-to-deliver (CTD), is the bond X that maximizes $Profit(X)$:

$$CTD \equiv \arg_X \max\{Profit(X)\}. \tag{3}$$

Although it might be a set of more than one bond, the CTD will be referred to as if it were just a single bond.

The following no-arbitrage conditions are proved in Appendix A:

No-arbitrage condition I *The delivery of any eligible bond X cannot produce a positive profit:*

$$Profit(X) \leq 0 \ , \quad for\ all\ X \tag{4}$$

Then, by (2), we have

$$FIP^X \leq S_T^X \ , \quad for\ all\ X \tag{5}$$

No-arbitrage condition II *Delivering the CTD generates zero profit:*

$$Profit(CTD) = 0 \tag{6}$$

Then, by (2), we have

$$FIP^{CTD} = S_T^{CTD} \tag{7}$$

From (6) and (3), the CTD is the only one that satisfies (4) with equality; therefore, delivering any alternative bond generates a loss. In other words, the CTD is the only bond for which the short receives a futures invoice price that equal its spot market price; for any alternative bond, the futures invoice price he receives is lower than its spot market price. It is clear, at this point, by looking at *(2),* that the criterion of minimum losses for delivering alternative bonds is equivalent to the criterion of closeness of the futures invoice prices to the corresponding spot market prices.

No-arbitrage condition II, which can be rewritten as $g\left(F_T, CTD\right) = S_T^{CTD}$ by substituting (1) into (7), will be used to derive the last settlement price F_T as a function of the price S_T^{CTD} of the CTD. Using this future-spot relation and no-arbitrage condition I, the rule to identify the CTD will be found.

4 The Conversion Factor System (CFS)

The CFS is currently the standard in all major exchanges of the world. It refers to *a particular functional form* $g\left(\cdot,\cdot\right)$ for (1). In the CFS, the futures invoice price for any eligible bond X is computed according to the following formula:

$$FIP^X = g(F_T, X) \equiv F_T \, CF^X \tag{8}$$

where FIP^X is the futures invoice price computed for a bond X, F_T is the last settlement price of the futures contract, and CF^X is a conversion factor (precomputed by the exchange) for bond X and for a futures contract expiring a T.

The conversion factor CF^X of bond X equals the would-be price per \$1 face value of bond X at the delivery day T when the remaining payments are discounted at a predetermined "reference rate":

$$CF^X = \sum_{t>T} C_t^X \left(1 + \frac{R}{2}\right)^{-2(t-T)} \tag{9}$$

where C_t^X represents the payment at time t per dollar of face value of bond X, t is the time of payment C_t^X, and R is the reference rate.[11]

The reference rate R is fixed by the specifications of the contract. In the case of the CBOT, it was 8% since the creation of the T-bond futures contract until it was changed to 6% beginning with March 2000 contracts.[12]

4.1 No-Arbitrage Relations at Expiration

This section reviews some no-arbitrage results for the CFS. It will facilitate the presentation of the new results for the TNBS in Section 5.2.

For the CFS, the absence of arbitrage opportunities implies:

$$F_T = \frac{S_T^{CTD}}{CF^{CTD}} = \min_X \left\{ \frac{S_T^X}{CF^X} \right\} \tag{10}$$

The first equality is obtained by specializing (8) for $X = CTD$ and substituting it into (7). It gives the relation between the last settlement price and the

[11] In reality, the accrued interests must me added to the last member of (8) and subtracted from the second member of (9). In order to make the exposition of the CFS easier to follow, I assume that T coincides with the beginning of a rental period of all eligible bonds; therefore, accrued interests are zero at T. Nevertheless, the results of the next section are still valid in general if we regard S_T^X as a clean price.

[12] It is instructive to compute the total *cash* proceeds for the *seller* during the whole life of the contract:

$$\begin{aligned} \textit{Total Proceeds for the Seller} &= \textit{Marking to Market Proceeds} + IP^X \\ &= -(F_T - F_0) + F_T \, CF^X \\ &= F_0 + F_T(CF^X - 1) \end{aligned}$$

where F_0 is the futures price originally contracted by the seller. If the short delivers a *6% coupon bond*, then $CF^X = 1$ because the coupon and the discount rate coincide; as a result, the short's total cash proceeds are equal to F_0, i.e. the originally contracted price. $F_T(CF^X - 1)$ is the correction intended to make the choice of the bond to deliver unimportant; it is positive when the coupon of the bond is higher than 6%, and negative otherwise.

spot price of the CTD at expiration. Given that the CTD is the bond that is chosen for delivery, it must be the one that determines F_T.

The ratio $\frac{S_T^X}{CF^X}$ is usually referred to as the futures-equivalent price corresponding to the spot price S_T^X of bond X.[13] The last equation in (10) tells us that the CTD is the eligible bond with lower futures-equivalent price: $CTD = \arg_X \min \left\{ \frac{S_T^X}{CF^X} \right\}$.

To prove the last equation in (10), as well as for future reference, we need a convenient expression of the "profit" of delivering any bond X. To specialize (2) for the CFS, substitute (8)

$$Profit(X) = F_T \, CF^X - S_T^X \tag{11}$$

Then, substitute $F_T = \frac{S_T^{CTD}}{CF^{CTD}}$ from (10) and rearrange:

$$Profit(X) = \left[\frac{S_T^{CTD}}{CF^{CTD}} - \frac{S_T^X}{CF^X} \right] CF^X \tag{12}$$

The last expression allows us to prove the last equality in (10): $\frac{S_T^{CTD}}{CF^{CTD}} = \min_X \left\{ \frac{S_T^X}{CF^X} \right\}$. By contradiction, assume that $\frac{S_T^{CTD}}{CF^{CTD}} > \frac{S_T^X}{CF^X}$ for some X; then, $Profit(X) > 0$, which implies the existence of an arbitrage opportunity by contradicting (4).

4.2 Rationale and Shortcomings of the CFS

The purpose of this section is to explain the rationale and to illustrate the shortcomings of the CFS in order to be able to compare it to the TNBS proposed in this paper.

According to Section 2, the objective of a futures invoice price function is that $Profit(X)$ not be too negative when the delivered bond X is different from the CTD. *Ideally*, CF^X and S_T^X should vary in the same proportion across X; in this way, $\frac{S_T^X}{CF^X}$ would be constant across X. As a result, (12) would be zero for all X and the delivery of any bond would be equivalent.

The CFS does part of the job. For example, the larger X's coupon, the larger both CF^X and S_T^X, thus producing some compensation in the ratio $\frac{S_T^X}{CF^X}$. However, in general, the CFS is far from making more or less equivalent the delivery of the any bond from the list of eligible issues, and, at times, it makes things even worse.

In order to illustrate the behavior of the CFS in a simple way, I take the case of a flat yield curve at T; then, the yield to maturity of all bonds is a constant y_T. Now, consider three different scenarios:

[13] The rationale for this terminology is that the futures-equivalent price $\frac{S_T^X}{CF^X}$ is the would-be last futures price F_T if X were the only eligible bond, as can be checked in (10).

(a) When the yield to maturity of all bonds is *equal* to the reference rate ($y_T = R$), then $S_T^X = CF^X$ for all X, and $\frac{S_T^X}{CF^X} = 1$ for all X. Therefore, the short is *indifferent* between delivering any bond by (12).

(b) When the yield of the bonds is *higher* than the reference rate ($y_T > R$), then $S_T^X < CF^X$ and $\frac{S_T^X}{CF^X} < 1$ for all X. This ratio will be lower, the higher the duration of bond X. (See Appendix B) Therefore, the *CTD* will tend to be a *high*-duration bond.

(c) When the yield of the bonds is *lower* than the reference rate ($y_T < R$), the *CTD* will tend to be a *low*-duration bond. (The argument is similar to that in (b).)

The system works optimally in the very special case (a); but the short suffers a loss if he delivers a bond other than the *CTD* in cases (b) or (c) because the differences in $\frac{S_T^X}{CF^X}$ across bonds will generate negative values in (12).[14] Nevertheless, even in the latter cases, the CFS can be of some help in reducing the aforesaid loss: *For equal maturities*, the bond with higher coupon will have a higher spot price and a higher conversion factor. As a result, there is some compensation that makes (11) less different across eligible bonds. *For a given coupon*, when this coupon is higher than the yield and the reference rate, longer bonds will have higher spot price and higher conversion factor; so, here we have certain compensation in (11). The compensations mentioned in this paragraph are less perfect, the bigger the difference between the yield and the reference rate.

The following is an example where the CFS produces the opposite to the intended effect. If the coupon (assumed fixed across bonds) is between the yield and the reference rate, $y_T < coupon < R$, then longer maturity bonds will have higher spot price and lower conversion factor. Therefore, the CFS increases the dispersion of (11) across eligible bonds. As the maximum value of $Profit(X)$ is zero, a bigger dispersion implies higher losses for delivering alternative bonds. A similar outcome obtains for $R < coupon < y_T$.

The imperfections of the CFS explained in scenarios (b) and (c), and in the last two paragraphs were the motivation to devise an alternative form $g(\cdot, \cdot)$ for the futures invoice price function (1), which is explained next.

5 The True Notional Bond System (TNBS)

In this section I propose a new procedure to determine the futures invoice price of each eligible bond at T, which will be referred to as the TNBS.

Notation. PV$[\cdot, \cdot]$ is a present value operator that gives the value at T of the bond identified by the first argument, where its cash flows are discounted at the rate in the second argument. As an example of the use of this operator, we have

$$S_T^X \equiv \text{PV}[X, y_T^X] \tag{13}$$

[14]These imperfections of the CF System have been noted by the literature. See, for example, Kane and Marcus (1986) or Schulte and Violi (2001).

were S_T^X is the full price of bond X, and y_T^X is its yield to maturity (YTM).

The following definition is an important building block of the futures invoice price determination procedure to be described in the next section.

Definition (Notional Bond). *The notional bond (NB) of a futures contract is an imaginary bond with the following characteristics:*

- *It is "issued" at the delivery date T of the futures contract.*
- *The coupon and tenor of this bond are fixed; they are part of the definition of the futures contract.*[15] *Then, we can determine exactly the amount and timing of the cash flows of the notional bond.*
- *The price of the notional bond at T is defined to be the last settlement price F_T of the futures contract: $S_T^{NB} \equiv F_T$.*

The yield to maturity of our NB at time T will be denoted by y_T^{NB}. Then, from the previous definition, we have the following identity

$$F_T \equiv \mathrm{PV}[NB, y_T^{NB}] \tag{14}$$

Some remarks about the definition of the notional bond: First, the notional bond is an *imaginary* bond. In principle, the futures Exchange could set any coupon and tenor in the specifications of the contract. However, it would be most natural to set them within the usual range of coupons and tenors of the eligible bonds. Second, to motivate the choice of F_T as the spot price of the NB at T, we can imagine that the notional bond is our underlying. Then, defining $S_T^{NB} \equiv F_T$ is a natural choice because, in any standard futures contract, the futures price converges to the spot price at expiration. Third, the Notional Bond defined above is, in fact, *just an element of the futures invoice price function* that I will propose. The real underlying assets are the bonds in the list of deliverable issues. Fourth, I have defined a *true* notional bond with a predetermined coupon and tenor. I remark "true" because some exchanges that use the CFS claim that they list futures contract on a "notional bond" with a coupon that is supposed to be the equal to the reference rate. However, the tenor of this "notional bond" is unspecified; therefore, there is not a true notional bond in the CFS. The CBOT is careful enough not to use this terminology.

5.1 Computing the Futures Invoice Price

Instead of stating directly the functional form $g(\cdot,\cdot)$ to be used in (1), I will describe a two-step procedure to compute the futures invoice price FIP^X. Later, I will collapse the two steps into one expression so as to formally specify $g(\cdot,\cdot)$.

To compute the futures invoice price FIP^X of bond X, the TNBS works as follows:

[15] For example, for the T-bond futures contracts whose underlying are bonds with at least 15 years to maturity, the notional bond could have a coupon of 6% and a tenor of 20 years.

1. Compute the internal rate of return of the notional bond from (14):

$$y_T^{NB} = \arg_y\{F_T = \text{PV}[NB, y]\} \tag{15}$$

2. Compute the futures invoice price of eligible bond X by discounting its cash flows at the rate y_T^{NB}:

$$FIP^X = \text{PV}\big[X, y_T^{NB}\big] \tag{16}$$

(While step 1. is common to all eligible bonds, step 2. is different for each.)

In summary, substituting (15) into (16), we get the proposed futures invoice price function:

$$FIP^X = g\left(F_T, X\right) \equiv \text{PV}\big[X\ ,\ \arg_y\{F_T = \text{PV}[NB, y]\}\big] \tag{17}$$

This futures invoice price function has the same arguments as the one used in the CFS: F_T and X. We could think of other functional forms including additional arguments, for example, futures prices on notes, or futures prices on bills. However, it is *un*likely that futures exchanges would be willing to accept a functional form that is much more complex to explain to the public. Therefore, I constrain the choice of functional forms to those whose only arguments are F_T and X. Later, I will suggest a refinement of the proposed functional form that also satisfies this constraint.

5.2 No-Arbitrage Relations at Expiration

In this section I show that, for the TNBS, the absence of arbitrage opportunities in a frictionless market imply

$$y_T^{NB} = y_T^{CTD} = \max_X \left\{y_T^X\right\} \tag{18}$$

For the sake of comparison to (10), we can reexpress (18) by applying PV[NB, .] to all three members of (18) and by using identity (14) in the first:

$$F_T = \text{PV}[NB, y_T^{CTD}] = \text{PV}[NB, \max_X \left\{y_T^X\right\}] \tag{19}$$

The first equality in (18) says that the YTM of the NB, as defined by (14), has to be equal to the YTM of the CTD.[16] To prove it, I start from equality (7), $FIP^{CTD} = S_T^{CTD}$, and use (16) and (13) to reexpress this equality as PV[CTD, y_T^{NB}] = PV[CTD, y_T^{CTD}], whence $y_T^{NB} = y_T^{CTD}$ trivially follows.

A key element in the analysis of the TNBS is the determination of the profit derived from delivering a particular bond X. As an intermediate step, we

[16] Note that $y_T^{NB} = y_T^{CTD}$ connects F_T with S_T^{CTD} making the first an increasing function of the second because F_T is decreasing in y_T^{NB}, and y_T^{CTD} is decreasing in S_T^{CTD}. Given that the CTD is the only eligible bond that is delivered, it must be the one that determines F_T.

substitute $y_T^{NB} = y_T^{CTD}$ into (16) to get

$$FIP^X = \mathrm{PV}\left[X, y_T^{CTD}\right] \tag{20}$$

Finally, substituting (20) and (13) into (2) we get

$$\begin{aligned} Profit(X) &= \mathrm{PV}[X, y_T^{CTD}] - \mathrm{PV}[X, y_T^X] && (21)\\ &\approx D^X S_T^X \left(y_T^X - y_T^{CTD}\right) && (22) \end{aligned}$$

where y_T^X and y_T^{CTD} are continuously compounded rates, and D^X is the duration of X.

Looking *only* at the YTM of the different eligible bonds, we do *not* know the *complete* ranking of eligible bonds according to the seller's preferences for delivery because, as we can see in (22), $D^X S_T^X$ also counts. However, we *can* say which is the *first* bond in the ranking (the *CTD*) by looking *only* at the YTM of the eligible bonds. The rule is that the *CTD* is the eligible bond with the highest YTM: $CTD = \arg_X \max\left\{y_T^X\right\}$. This rule derives from the second equality of (18), $y_T^{CTD} = \max_X \left\{y_T^X\right\}$, which I prove next. By contradiction, assume that $y_T^{CTD} < y_T^X$ for some X; then (21) and (22) are positive, which implies the existence of an arbitrage opportunity by contradicting (4).

5.3 Rationale of the TNBS and Comparison with the CFS

Bonds of similar maturities have similar yields. Therefore, if the eligible bonds belong to a convenient range of maturities, their yields will be similar. In this case, (21) and (22) will be small, and delivering a bond different from the *CTD* will generate only a small loss. This was precisely the metric proposed in Section 2 to judge alternative forms $g(\cdot, \cdot)$ for the futures invoice price function (1), hence the rationale of the TNBS.

While the range of maturities needs to be narrow for futures contracts on short term notes, it can be wide for long term bonds because the latter's yields tend to differ much less. The note and bond futures contracts at the CBOT follow exactly this pattern; therefore, the futures invoice price function proposed here fits the current structure of the contracts at the CBOT.

Another way to see the rationale of the TNBS: given that the yields of the eligible bonds are similar, the futures invoice price (20) received by the short for delivering any bond will be close to its market value.

Now, I compare the TNBS with the CFS. Let us say that a futures invoice price determination system *works perfectly* when $\forall X : Profit(X) = 0$. (This is consistent with Section 2.) Assume that the yield curve is *flat*. While the CFS works perfectly only when the level of the yield is equal to the 6% reference rate, the TNBS works perfectly for *any* level of yield by (21) or (22). Thus, the world in which the CFS works perfectly is much more implausible than the one in which the TNBS does. It is somewhat ironic that the CFS only works perfectly in a world in which the yields are fixed, in which case a bond futures contract would be useless. The empirical part of this paper shows that the proposed system is superior to the CFS also for real-world non-flat yield curves.

Another disadvantage of the CFS is that, when yields become too different from the reference rate, the Exchange has to change the reference rate. The CBOT was recently forced to change the reference rate from 8% to 6% to diminish the losses of delivering a bond different from the cheapest-to-deliver. In contrast, no change in the computation of the futures invoice price is ever needed in the TNBS. Although this may seem a minor consideration, in practice it is not: the CBOT had to widely publicize the year 2000 change in the reference rate and its implications in order to avoid confusions, and needed the approval by the Commodity Futures Trading Commission.

The remarks of the two previous paragraphs are particularly important in emerging countries, where the interest rates include ever-changing credit spreads, which make interest rates more volatile than in developed economies. Therefore, the yields can easily depart very substantially from the reference rate.

A word about hedging. In the CFS, hedgers usually compute the sensitivity of the futures price to changes in the interest rate assuming that a particular bond will turn out to be the *CTD*. When the *CTD* is still uncertain, hedging becomes problematic and may require frequent rebalancing of the hedger's portfolio. As an example of this, cases (b) and (c) of Section 4.2 imply that, under a flat yield curve, a small increase in the yield from being lower to being higher than the reference rate R makes the *CTD* switch dramatically from low duration bonds to high duration bonds. In contrast, under the TNBS, hedgers can design the typical hedging strategies quite easily: they only need to remember that the futures price is the present value of the notional bond and, therefore, they can act as if the underlying of the futures contract were the notional bond. The most common duration-based strategies assume that the yield curve suffers only parallel shifts. Under this assumption, the change in the yield y^{NB} of the notional bond is equal to the change in the rest of the yields because y^{NB} is just their maximum.

CBOT (2000) considers a frequently asked question: "How do you determine the yield of a Treasury futures contract based on its price?" This paper concludes that, in the CFS, futures contracts do not have a yield.[17] In contrast, in the TNBS there is a clear yield, namely, the yield of the notional bond. This yield can be compared to the yields of Treasury bonds, knowing that, at expiration, y_T^{NB} will equal the highest among the yields of all eligible bonds. This feature adds to the contract's appeal to speculators and the public in general.

Finally, we can consider the effect of the futures market on the spot market: the anomalies at the far end of the spot yield curve will tend to be lower with the TNBS than with the CFS. Even with an overall upward sloping yield curve, it is

[17] If we knew in advance at time t which bond will be the *CTD* at T, then we would be able to say that the yield of the futures contract at t is the yield that the *CTD* would have at T if its price at this time were $F_t\, CF^{CTD}$ (known at t). However, the *CTD* is not known in advance. Furthermore, even at T, the yield of the futures contract may not be unique if two or more bonds are *CTD* at the same time.

common to see a negative slope in the far end of the yield curve. This is usually attributed to the fact that the longest bonds have been issued more recently and therefore still remain more liquid. The CFS can stress this anomalous negative slope in the far end of the yield curve: when the yields are higher than the reference rate, the highest-duration bond tends to be cheaper to deliver, pushing its price up and its yield down. On the contrary, the TNBS will tend to flatten the far end of the yield curve because the *CTD* will be the bond *with higher YTM*. Then, an increased demand of the *CTD* in anticipation for delivery will drive its price up and its YTM down.

6 Empirical Comparison of the Two Systems

This section empirically compares the CFS and the TNBS according to the losses generated by delivering bonds different from the cheapest-to-deliver. The rationale for the chosen criterion of comparison was explained in Section 2. The focus is on the Treasury-bond futures contract traded on the CBOT. The eligible bonds are long-term U.S. Treasury bonds that mature in at least 15 year's time from the expiration of the corresponding futures contract.[18] First, for reasons that I explain later, results are presented using simulated bonds that are valued from an estimated yield curve model. Then, as a robustness check, I present results using actual Treasury bond prices.

The Data. The sample contains US Treasury bonds and notes included in the CRSP's Monthly US Treasury Database from February 1985 to December 2001. Only taxable, non-flower, non-callable bonds and notes were used.

Yield Curve Model. The Nelson and Siegel (1987) yield curve model was fitted to all the data of the sample. This model assumes that the functional form of the instantaneous forward rate at t as a function of maturity m is

$$f_t^m = \beta_{1,t} + \beta_{2,t} e^{-\frac{m}{\delta_1}} + \beta_{3,t} \frac{m}{\delta_2} e^{-\frac{m}{\delta_2}} \tag{23}$$

whence we can get the functional form of the zero-coupon yields at t as a function of maturity m:

$$y_t^m = \beta_{1,t} + \beta_{2,t} \frac{\delta_1}{m} \left[1 - e^{-\frac{m}{\delta_1}}\right] + \beta_{3,t} \left[\frac{\delta_2}{m}\left(1 - e^{-\frac{m}{\delta_2}}\right) - e^{-\frac{m}{\delta_2}}\right]$$

δ_1 and δ_2 were constrained to be the same for all the dates of the sample. Then the evolution of the linear parameters $\left\{\beta_{1,t}, \beta_{2,t}, \beta_{3,t}\right\}_t$ characterizes the evolution of the yield curve through time. Nelson and Siegel (1987) and Diebold and Li (2002) have noticed that very little goodness-of-fit is sacrificed when δ_1 and δ_2 are constrained to be constant through time, even when they force $\delta_1 = \delta_2$.

[18] In addition, eligible bonds should not be callable for at least 15 years. I will ignore the possibility of delivering callable bonds because, at present, the US Treasury is not issuing them any more and there are not any eligible callable bonds left for the T-bond futures contract.

6.1 Results

Although there only exist futures contracts expiring every three months, I simulated the behavior of futures expiring every single month, thus taking advantage of all the data points in the sample.

Using the estimated yield curves, I simulated the prices of bonds of different maturities for each month of the sample. The chosen maturities range from 15 to 30 years with a step of half a year; thus I work with 31 bonds at each expiration, which closely approximates the number of deliverable issues in the CBOT's Treasury Bond Futures Contract. The coupon of the simulated bonds is chosen to be 8%; the results are robust to the choice of coupon.

The simulated prices were the input to simulate, at each time, the YTM of the different bonds, the yield y^{NB} of the notional bond using (18), the futures invoice prices of the bonds under the TNBS using (16), and the corresponding losses using (2). The same simulated prices were the input to compute, for each month in the sample, the futures price under the CFS using (10), and the losses from delivering the different bonds using (11).

Recall that "alternative bonds" denote the eligible bonds that are not the *CTD*. Let $l_{t,i}$ be the loss under the CFS derived from the seller's delivering, at date t, \$1 face value of the alternative bond i. For each date t, the losses $\{l_{t,i}\}_{i=1}^{30}$ for the 30 alternative bonds were sorted in ascending order to obtain $\{L_{t,\#}\}_{\#=1}^{30}$ such that $L_{t,\#} \leq L_{t,\ \#+1}$. This sorting determines the ranking # of the alternatives to the *CTD* according to the seller's preferences. Now suppose that we pick the loss of the $\#th$ ranked bond for each of the n months in the sample: $\{L_{t,\#}\}_{t=1}^{n}$. The average $\bar{L}_{\#} = \frac{1}{n}\sum_{t=1}^{n} L_{t,\#}$ of those values is shown by the corresponding round marker of Figure 1. The triangular markers show the results of a parallel ordering and the same computations for the losses corresponding to the TNBS. From the procedure described above, the markers of Figure 1 show the losses of the cheaper bonds to deliver to the left and the more expensive to deliver to the right.

The higher horizontal line in Figure 1 indicates the average of the round markers, $\overline{\overline{L}} = \frac{1}{m}\sum_{\#=1}^{m} \bar{L}_{\#}$, i.e. the average loss of delivering each and every bond simulated at each and every month of the sample period for the CFS. The lower horizontal line shows the parallel computation for the TNBS. As we can see, the average loss in the CFS is more than twice the one of the TNBS. However, this figure focuses our attention on the less important losses, the ones that will never occur, because the higher differences are those corresponding to the less preferred bonds.

A better picture of the relative losses generated by the two systems can be obtained by dividing the height of each marker by that of the marker located in the same vertical line corresponding to the TNBS. The resulting ratios are the markers of Figure 2. In this way, for each #, the average losses corresponding to each of the two systems have been standardized so as to have 1 for the TNBS.

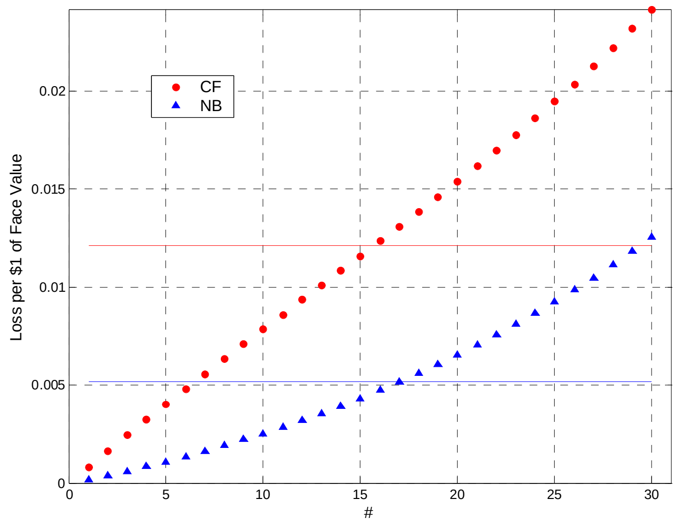


Figure 1: *Average loss per \$1 of face value from delivering the #th best alternative to the CTD.*

The horizontal line is just the average of the markers corresponding to the CFS.

In Figure 2, we can clearly see the most important comparison, the one between the losses in the left side of the graph: those in the CFS are of the order of four times those of the TNBS. Table 1 displays the results discussed so far.

6.2 Why use simulated prices?

The reason for using simulated prices is that actual prices of the *CTD* are usually overvalued because of the demand of short future traders and, possibly, of market manipulators. This overpricing of the *CTD* has two effects:

1. By the first equation in (10), the futures contract becomes overpriced; therefore, the short must make excess payments of variation margins.

2. The overpricing of F_T increases the futures invoice price received by the short, thus reducing the losses of delivering alternative bonds.[19]

As the CFS is the futures invoice price determination procedure used in the real world, the bonds that are the *CTD according to this system* are usually overpriced, even long before the expiration of the futures contract. Therefore, according to 2., using actual prices would reduce the losses for the CFS reported in the previous figures. However, this reduction of losses is not something to

[19]The effect of an overpricing of the *CTD* on the losses of delivering alternative bonds can be seen in (12).

#	CF	True NB	Ratio
1	0.00083	0.00019	4.29
2	0.00165	0.00040	4.11
3	0.00246	0.00062	3.95
4	0.00325	0.00086	3.80
5	0.00404	0.00110	3.67
6	0.00481	0.00136	3.54
7	0.00558	0.00163	3.42
8	0.00634	0.00192	3.30
9	0.00710	0.00222	3.20
10	0.00785	0.00254	3.10
11	0.00860	0.00287	3.00
12	0.00935	0.00321	2.91
13	0.01009	0.00357	2.83
14	0.01084	0.00394	2.75
15	0.01159	0.00434	2.67
16	0.01234	0.00475	2.60
17	0.01309	0.00517	2.53
18	0.01384	0.00562	2.46
19	0.01461	0.00608	2.40
20	0.01539	0.00657	2.34
21	0.01617	0.00706	2.29
22	0.01697	0.00758	2.24
23	0.01779	0.00812	2.19
24	0.01862	0.00868	2.15
25	0.01947	0.00926	2.10
26	0.02035	0.00987	2.06
27	0.02126	0.01049	2.03
28	0.02221	0.01116	1.99
29	0.02319	0.01186	1.96
30	0.02419	0.01258	1.92
Average	0.01213	0.00519	2.79

Table 1: *Comparison between the losses (per dollar of face value) generated by the CF and the True NB Systems.* The last column is the ratio of the previous two.

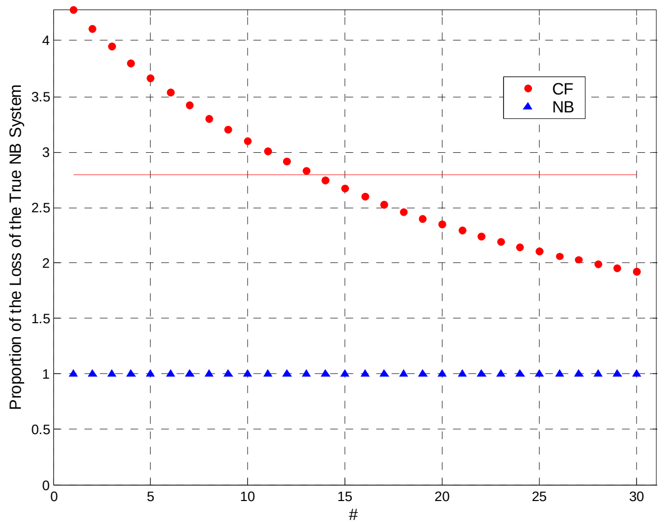


Figure 2: *Standardized average loss from delivering the #th best alternative to the CTD.* Measured as a proportion of the corresponding average loss in the NB System.

celebrate because it is compensated by the losses in the futures explained in 1.; in fact, the latter are much more harmful because they affect the short even when he delivers the *CTD*. Therefore, it is preferable that the overpricing of the *CTD* does not reduce the reported losses. To avoid taking this overpricing into account, I used simulated prices obtained from a parsimonious representation of the yield curve so as to simulate a spot market not distorted by pressures from the futures market.

In addition, there may be other specialness factors that are not caused by the futures contracts. Given that no attempt is made to consider them under either the CFS or the TNBS, it is desirable that they do not play any role in the comparison of the two systems. The use of a smooth representation of the yield curve is useful to wash them away.

Finally, using simulated prices fills the big gaps in maturities observed especially in the first half of the sample. As these gaps are not present any more in today's market, it would be of no interest to let them influence the results.

In spite of the previous comments, I show the results using *actual prices* as a robustness check. I restrict the experiment to data from February 1991 onward, the reason being that, from that month on, there are at least 16 bonds per month allowing me to compute at least 15 losses for all months; also, since that month there are no big differences between the maturity of one bond and that of the closest-in-terms-of-maturity bond. Given that in recent times we do not see any more those big maturity "holes", it would be of no interest to include the effect of such holes in the results. This is important given that the first

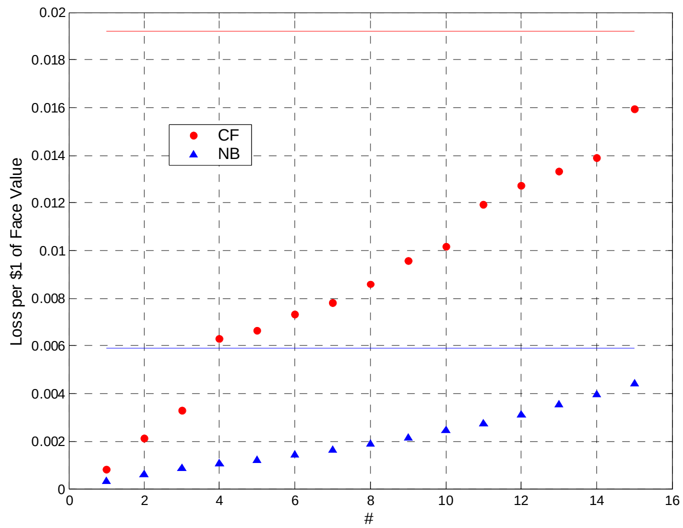


Figure 3: *Average loss per \$1 of face value from delivering the #th best alternative to the CTD.* (Using observed prices; computed to check robustness.)

preferred alternatives to the *CTD* are usually bonds of similar maturities.

Figures 3 and 4 are the equivalent of Figures 1 and 2 except for the use of actual prices to compute losses. The horizontal lines of Figures 3 and 4 were computed as an average of the losses of delivering each of the bonds available at each date, not only the first 15 alternative bonds, whose average losses are shown by the markers; that is why they seem out of line in Figure 3.

The first three round markers of Figure 4 have a height of the order of 3, somewhat less than the ones of Figure 2, possibly reflecting the overpricing of the cheaper bonds to deliver. However the CFS still compares very unfavorably to the TNBS.

7 Possible Extensions

I showed in Sections 4.2 and 5.3 that the TNBS works perfectly under flat yield curves. In that situation, any bond can be delivered without any loss. In this section, I suggest some extensions that could be used to fine tune the TNBS. Their purpose is to deal better with non-flat yield curves while keeping the futures invoice price as a function of only F_T and X. All of them have a common structure, which I denote the *Refined Version* of the TNBS. This Refined Version would work as follows. The Exchange fixes a spread s^X for each eligible bond X of the list of deliverable issues of a new futures expiration *before this expiration is listed*. In computing the futures invoice price of each bond X, the prespecified spread s^X will be added to y_T^{NB}, and the sum of the

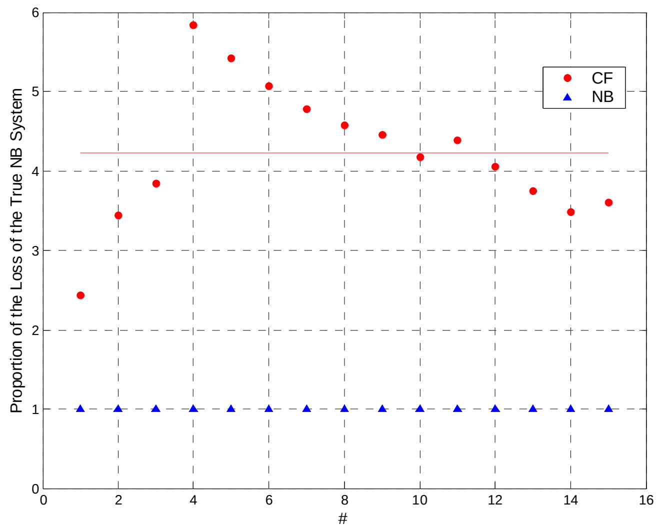


Figure 4: *Standardized average loss from delivering the #th best alternative to the CTD.* Measured as a proportion of the corresponding average loss in the NB System. (Using observed prices; computed to check robustness.)

two will be used to discount the cash flows of X:

$$FIP^X = \mathrm{PV}\big[X \ , \ y_T^{NB} + s^X\big] \tag{24}$$

This expression replaces (16) while y_T^{NB} is still computed as in (15).

In order to emphasize the timing of the determination of the spread, let τ be the time the spreads $s^X, \forall X$ are fixed by the Exchange for a particular futures expiration and υ the time this expiration begins to be traded, then we will require:

$$\tau < \upsilon < T$$

We need $\tau < \upsilon$ so that the contract be perfectly defined before trading begins.

If, at time τ, the expected shape of the yield curve for time T is different from flat, then the introduction of well chosen spreads may reduce the expected losses from delivering alternative bonds.

Condition for Indifference in the Delivery of Eligible Bonds. Using arguments similar to the ones of Section 5.2, the following can be proved. If, at time τ, the Exchange fixed spreads such that it turns out later at T that

$$s^X = y_T^X - y_T^Z \ , \quad for \ all \ X$$

where Z is an arbitrary (possibly imaginary) bond, then the delivery of any eligible bond is indifferent for the short. In other words, when the shape of the term structure of spreads s^X is equal to the shape of the term structure of

YTMs y_T^X at T of the eligible bonds, then indifference in delivery is achieved.

How the Exchange can fix the spreads at τ. Because of the previous indifference condition, an simple rule to fix the spreads would be

$$s^X = \mathrm{E}\big[y_T^X - y_T^Z\big] \ , \quad for\ all\ X$$

The better the Exchange can forecast the differences in YTMs, the less important will be for the short to deliver any of the eligible bond.

It is often claimed that more than 75% of the movements of the yield curve are parallel shifts. This fact should make the forecast of $y_T^X - y_T^Z$ more precise because this difference is almost[20] independent of the level of the yield curve.

A crude forecast of $y_T^X - y_T^Z$ is $y_\tau^X - y_\tau^Z$. A more sophisticate and general procedure is to forecast the zero-coupon yield curve to be observed at T, and then to compute forecasted prices and the corresponding forecasts for the YTMs. The yield-curve forecast can be performed in an infinity of ways; see for example Duffee (2002) and Diebold and Li (2002).

If the Exchange is reluctant to fix the spreads based on a new estimation of the zero-coupon yield curve every time a new expiration is about to be listed, it may consider using its unconditional average shape. Although the latter is a poor forecast, recall that the level is not important, only the shape is. Using this forecast might still be better that assuming that the yield curve will be flat. The advantage of this procedure is that, once the unconditional estimates have been obtained, the estimates can be frozen and, therefore, the rule for the determination of the spread is automatic: there is not need of further estimations each time a new futures expiration is about to be listed. Then, this rule can be incorporated in the definition of the contract on a permanent basis.

On the MATIF, there is a bond futures contract that includes both French and German government bonds in the list of deliverable issues. Apart from the usual problems of the CFS, this contract has the drawback that the *CTD* is almost always a French bond. Then the contract can hardly benefit from the addition of the German bonds to the basket of deliverable issues. (See Schulte and Violi (2001)) This could be handled by forecasting two different yield curves, one for each issuer, and using the corresponding one in determining the spread of each bond. If the Exchange preferred a futures invoice price determination system closer to the simple version of the TNBS, then s_X could be 0 for any bond X corresponding to one of the countries, and a constant $s_{X'}$ for any X' corresponding to the other country.

In addition to the general shape of the yield curve, we could take into account the idiosyncratic part of the YTM of each bond due, for example, to liquidity issues or to differences in seniority.

[20] A parallel shift of Δy in the yield curve produces a change in the YTMs of the bonds only *approximately* equal to Δy, as it has been noted by Ingersoll, Skelton and Weil (1978)).

8 Conclusion

In this paper I have proposed a new method, the TNBS, for determining the invoice price in futures contracts where the seller chooses one bond to deliver out of a list of deliverable issues. Assuming no-arbitrage, I have derived the following theoretical results for the proposed system: the equilibrium futures price at expiration, its relation to the price of the *CTD*, the equilibrium determination of the cheapest-to-deliver, and the equilibrium futures invoice price of all eligible bonds. Finally, after explaining the rationale of the TNBS, I have empirically shown that it strictly dominates the CFS, the current standard in the industry.

The empirical part of this paper shows that the TNBS dramatically reduces the losses derived from delivering a bond different from the cheapest-to-deliver. As a result, the TNBS gives the seller an economically more real possibility of delivering alternative bonds, provides a tighter bound to the overpricing of the cheapest-to-deliver and of the futures contract, and lessens the incentives to deliberately provoke shortages of the cheapest-to-deliver.

The *Refined Version* of the TNBS may reduce even further the losses of delivering alternative bonds. In addition, it allows a reasonable design for futures contracts whose list of deliverable securities comprises bonds issued by different countries. Finally, it permits the inclusion of bonds with different liquidity and seniority. The last point may be important for bonds issued by less developed countries.

The TNBS is conceptually clear. One can think of the futures contract as if the underlying were a single bond: the notional bond. For practical purposes like hedging or speculating on the level of the yield curve, the error of this only approximate characterization can be ignored. In addition, a yield can be meaningfully and straightforwardly ascribed to the futures price.

In future work I plan to 1) address the valuation of bond futures under the TNBS, 2) quantify the advantages of the TNBS over the CFS in the use of bond futures for risk management, and 3) treat in detail the Refined Version of the TNBS.

Appendix A

Proof of no-arbitrage condition I

If $Profit(X) > 0$, arbitrageurs can short the futures contract at T, choose to deliver bond X and make a profit. The reverse arbitrage cannot be done for *any* X because the long cannot choose which bond the short will deliver.

Proof of no-arbitrage condition II

For $X = CTD$, (4) implies $Profit(CTD) \leq 0$. Now I show that also $Profit(CTD) \geq 0$. If $Profit(CTD) < 0$, arbitrageurs will buy futures at T, and the worst that can happen to them in this zero-sum game is that the short delivers the *CTD*.

Then the arbitrageur will make at least a profit of $-Profit(CTD) > 0$. (Remember that $Profit(X)$ is the profit of the seller; to get the one of the buyer we have to change its sign.)

Appendix B

The influence of the level of yields on the determination of the cheapest-to-deliver in the CFS

This Appendix shows that, for a flat yield curve, when the yield of the bonds is *higher* than the reference rate, then the *CTD* will tend to be a *high*-duration bond, and that, on the contrary, when the yield is *lower* than the reference rate, then the *CTD* will tend to be a *low*-duration bond.

The argument will make use of four elements. First, y_T will be the *continuously compounded* yield to maturity of all bonds at time T:

$$S_T^X = \sum_{t>T} C_t^X \, e^{-y_T(t-T)} \ , \quad for \ all \ X \tag{25}$$

Second, the duration of a bond X is equal to

$$D^X = -\frac{dS_T^X}{dy_T} \frac{1}{S_T^X} = -\frac{d \log S_T^X}{dy_T} \tag{26}$$

Third, R, the reference rate, will be used as a continuously compounded yield, such that (9) becomes

$$CF^X = \sum_{t>T} C_t^X e^{-R(t-T)} \tag{27}$$

Fourth, recall the rule that identifies the *CTD*:

$$CTD = \arg_X \min \left\{ \frac{S_T^X}{CF^X} \right\} \tag{28}$$

When the yield of all bonds, a constant y_T, is *higher* than the reference rate ($y_T > R$), then $S_T^X < CF^X$ for all X. The higher the duration of X, the greater tends to be the absolute value of the difference between $\log S_T^X$ and $\log CF^X$ by (26), the lower $\log \frac{S_T^X}{CF^X}$, and the lower the futures-equivalent price $\frac{S_T^X}{CF^X}$. Then, by (28), the *CTD* will tend to be a *high*-duration bond.

A similar argument shows that, on the contrary, when the yield is *lower* than the reference rate ($y_T < R$), the *CTD* will tend to be a *low*-duration bond.

References

Boyle, P. P.: 1989, The quality option and timing option in futures contracts, *Journal of Finance* **44**, 101–113.

Carr, P. and Chen, R.-R.: 1997, Valuing bond futures and the quality option. Working Paper, Rutgers University.

CBOT: 2000, Futures and yields in a 6% world. Chicago Board of Trade Publication.

Cox, J. C., Ingersoll, J. E. and Ross, S. A.: 1981, The relation between forward prices and futures prices, *Journal of Financial Economics* **9**, 321–346.

Diebold, F. X. and Li, C.: 2002, Forecasting the term structure of government bond yields. Center for Financial Institutions Working Paper 02-34, Wharton School Center for Financial Institutions, University of Pennsylvania.

Duffee, G. R.: 2002, Term premia and interest rate forecasts in affine models, *Journal of Finance* **57**, 405–443.

Gay, G. D. and Manaster, S.: 1984, The quality option implicit in futures contracts, *Journal of Financial Economics* **13**, 353–370.

Hegde, S. P.: 1988, An empirical analysis of implicit delivery options in the Treasury bond futures contract, *Journal of Banking and Finance* **12**, 469–492.

Hegde, S. P.: 1990, An ex post valuation of the quality option implicit in the Treasury bond futures contract, *Journal of Banking and Finance* **14**, 741–760.

Hemler, M. L.: 1990, The quality delivery option in Treasury bond futures contracts, *Journal of Finance* **45**, 1565–1586.

Ingersoll, J. E., Skelton, J. and Weil, R. L.: 1978, Duration forty years later, *Journal of Financial and Quantitative Analysis* **13**, 627–650.

Johnston, E. T. and McConnell, J. J.: 1989, Requiem for a market: An analysis of the rise and fall of a financial futures contracts, *Review of Financial Studies* **2**, 1–23.

Jordan, B. D. and Kuipers, D. R.: 1997, Negative option values are possible: The impact of Treasury bond futures on the cash U.S. Treasury market, *Journal of Financial Economics* **46**, 67–102.

Kamara, A. and Siegel, A. F.: 1987, Optimal hedging in futures markets with multiple delivery specifications, *Journal of Finance* **42**, 1007–1021.

Kane, A. and Marcus, A. J.: 1984, Conversion factor risk and hedging in the Treasury-bond futures market, *Journal of Futures Markets* **4**, 55–64.

Kane, A. and Marcus, A. J.: 1986, The quality option in the Treasury bond futures market: An empirical assessment, *Journal of Futures Markets* **6**, 231–248.

Livingston, M.: 1987, The delivery option on forward contracts, *Journal of Financial and Quantitative Analysis* **22**, 79–87.

Magdon-Ismail, M., Atiya, A. F. and Abu-Mostafa, Y. S.: 2000, Pricing the quality option for the bond futures contract in a multifactor Vasicek framework. Proceedings of the 16th IMACS World Congress, Lausanne, Switzerland, August 2000.

Merrick, J. J., Naik, N. Y. and Yadav, P. K.: 2002, Strategic trading behavior and price distortion in a manipulated market: Anatomy of a squeeze. Forthcoming in Journal of Financial Economics.

Nelson, C. R. and Siegel, A. F.: 1987, Parsimonious modeling of yield curves, *Journal of Business* **60**, 473–489.

Schulte, W. and Violi, R.: 2001, Interactions between cash and derivatives bond markets: Some evidence for the Euro area, *BIS Papers* **5**, 67–112.

Vidal Nunes, J. P. and Ferreira de Oliveira, L. A.: 2003, Quasi-analytical multifactor valuation of the treasury bond futures with an embedded quality option. Proceedings of the 10th Annual Meeting of the German Finance Association.

Yu, S.-W.: 1999, Delivery options and hedging effectiveness, *Advances in Pacific Basin Financial Markets* **5**.

The Determinants of Credit Default Swap Premia

Jan Ericsson, Kris Jacobs, and Rodolfo Oviedo
Faculty of Management, McGill University*

First Version: May 2004
This Revision: January 2005

Abstract

Using a new dataset of bid and offer quotes for credit default swaps, we investigate the relationship between theoretical determinants of default risk and actual market premia using linear regression. These theoretical determinants are firm leverage, volatility and the riskless interest rate. We find that estimated coefficients for these variables are consistent with theory and that the estimates are highly significant both statistically and economically. The explanatory power of the theoretical variables for levels of default swap premia is approximately 60%. The explanatory power for the differences in the premia is approximately 23%. Volatility and leverage by themselves also have substantial explanatory power for credit default swap premia. A principal component analysis of the residuals and the premia shows that there is only weak evidence for a residual common factor and also suggests that the theoretical variables explain a significant amount of the variation in the data. We therefore conclude that leverage, volatility and the riskfree rate are important determinants of credit default swap premia, as predicted by theory.

JEL Classification: G12

Keywords: credit default swap; credit risk; structural models; leverage; volatility.

---

*Ericsson is a research affiliate of SIFR. Jacobs is affiliated with CIRANO and CIREQ. This paper has benefited from the comments of David Chapman, Yigal Newman, Olivier Renault and Jun Yang, as well as seminar participants at HKUST, the European Investment Review 2004 meetings and the Bank of Canada Fixed Income Workshop. Ericsson would like to thank FCAR and IFM^2 for financial support and SIFR for making the default swap data available to us. Jacobs is grateful for financial support from FCAR, IFM^2 and SSHRC. Oviedo was supported in part by a grant of IFM^2. Any remaining inadequacies are ours alone.
Correspondence to: Jan Ericsson, Faculty of Management, McGill University, 1001 Sherbrooke Street West, Montreal, Canada H3A 1G5; Tel: (514) 398-3186; Fax: (514) 398-3876; E-mail: jan.ericsson@mcgill.ca.

1 Introduction

A credit derivative is a contingent claim that allows the trading of default risk separately from other sources of uncertainty. From being a fledgling market in the mid nineties, credit derivative markets have grown tremendously over the last few years. The market exceeded 2 trillion dollars in outstanding notional principal in 2002, and it is expected to double in size by the end of 2004. The most standard contract is the single-name credit default swap (CDS) which accounts for roughly half of the trading activity.[1] This instrument is essentially an insurance contract against the default of an underlying entity. Compensation is paid if a credit event occurs while in return the buyer of protection makes regular payments based on the *swap premium*.

Little empirical work has been done on credit derivative markets.[2] Notable exceptions include Berndt, Douglas, Duffie, Ferguson, and Schranz (2004), Houweling and Vorst (2005), Hull, Predescu, and White (2004) and Longstaff, Mithal, and Neis (2004). Berndt, Douglas, Duffie, Ferguson, and Schranz (2004) use CDS data, among other things, to study default risk premia. Houweling and Vorst (2005) implement a set of simple reduced form models on market CDS quotes and corporate bond quotes. The paper focuses on the pricing performance of the model and the choice of benchmark yield curve. Hull, Predescu, and White (2004) analyze the impact of rating announcements on the pricing of CDSs. Longstaff, Mithal, and Neis (2004) and Blanco, Brennan, and Marsh (2003) study the relative pricing of corporate bonds and default swaps.

In the last decade, a more substantial body of empirical work has emerged on other credit sensitive instruments, in particular corporate bonds. This work can be categorized according to the theoretical framework it relies on. One popular approach is to use what are known as reduced form models.[3] These models exogenously postulate the dynamics of default probabilities and use market data to recover the parameters needed to value credit sensitive

[1] These statistics and forecasts are based on publications by the British Bankers' Association . A very similar picture emerges from our dataset. Although it includes some transactions that date back to 1995, the number of quotes is negligible until the turn of the century. Subsequently the market experienced explosive growth (see Figure 1).

[2] Theoretical work includes Das (1995), Hull and White (2000) and Das and Sundaram (1998).

[3] See Jarrow and Turnbull (1995) and Duffie and Singleton (1999) for early work on this topic. Useful surveys can be found in Lando (1997) and Duffie and Singleton (2003).

claims.[4] While these models have been shown to be versatile in practical applications, they remain relatively silent on the theoretical determinants of the prices of defaultable securities.

An alternative approach, commonly referred to as the structural approach, is to rely on models that have evolved following Black and Scholes (1973) and Merton (1974). This approach links the prices of credit risky instruments directly to the economic determinants of financial distress and loss given default.[5] In particular, these models imply that the main determinants of the likelihood and severity of default are financial leverage, volatility and the risk free term structure. These models have been plagued by poor performance in empirical studies.[6] Perhaps as a result of the difficulty of implementing structural models in practice, a more direct approach was taken by Collin-Dufresne, Goldstein, and Martin (2001) (CGM), who use the structural approach to identify the theoretical determinants of corporate bond credit spreads. These variables are then used as explanatory variables in regressions for changes in corporate credit spreads, rather than inputs to a particular structural model. CGM conclude that the explanatory power of the theoretical variables is modest, and that a significant part of the residuals are driven by a common systematic factor which is not captured by the theoretical variables. Campbell and Taksler (2003) (CT) perform a similar analysis but use regressions for levels of the corporate bond spread. They conclude that firm specific equity volatility is an important determinant of the corporate bond spread and that the economic effects of volatility are large. Cremers, Driessen, Maenhout, and Weinbaum (2004) (CDMW) confirm this result, and argue that option-based volatility contains information useful for this type of analysis that is different from historical volatility.

Our study is intimately related to these papers. Although our focus is also on credit risk, an important distinction is that we study very different data – default swap premia rather than corporate bond yield spreads. Using default swaps rather than bonds has at least two

[4]Empirical papers using reduced form models to value credit risky bonds include Bakshi, Madan, and Zhang (2001), Driessen (2004), Duffee (1999), Duffie and Lando (2000), Duffie, Pedersen, and Singleton (2003) and Elton, Gruber, Agrawal, and Mann (2001).

[5]Important examples include Black and Cox (1976), Collin-Dufresne and Goldstein (2001), Geske (1977), Kim, Ramaswamy, and Sundaresan (1993), Leland (1994), Leland and Toft (1996), Longstaff and Schwartz (1995) and Nielsen, Saá-Requejo, and Santa-Clara (1993).

[6]See in particular Jones, Mason, and Rosenfeld (1984), Jones, Mason, and Rosenfeld (1985), Lyden and Saranati (2000) and Ogden (1987). More recently Eom, Helwege, and Huang (2004) have documented the difficulty of implementing these models.

important advantages.

First, default swap premia, while economically comparable to bond yield spreads, do not require the specification of a benchmark risk free yield curve – they are already "spreads". Thus we avoid any added noise arising from a misspecified model of the risk free yield curve. The choice of the risk free yield curve includes the choice of a reference risk free asset, which can be problematic (see Houweling and Vorst (2005)), but also the choice of a framework to remove coupon effects.

Second, default swap premia may reflect changes in credit risk more accurately and quickly than corporate bond yield spreads. Blanco, Brennan, and Marsh (2003) provide evidence that changes in the credit quality of the underlying name are likely to be reflected more quickly in the swap premium than in the bond yield spread. Also, if there are other important non-default components in bond spreads, their variation will obscure the impact of changes in credit quality.[7]

Like CGM, CT and CDMW, we carry out linear regression analysis on the relationship between default swap premia and key variables suggested by economic theory. Our benchmark results focus on financial leverage, firm specific volatility and the risk free rate. We run regressions on changes in premia as well as for the levels of the premia. We find that the estimated coefficients for the three variables are consistent with theory and that the estimates are highly significant both statistically and economically. The size of the effects is intuitively plausible. This is true both for regressions in levels and differences. Interestingly, we find a negative correlation between CDS premia and the risk free rate. A similar correlation has been documented for bond yield spreads by Longstaff and Schwartz (1995) and Duffee (1998). Presently, no consensus prevails as to the economic reasoning behind this stylized fact. Our results are consistent with the implication of structural models that an increase in the risk free rate will decrease risk-adjusted default probabilities.

The amount of the variation in swap premia explained by the difference regressions is higher than in existing work on corporate bond spreads. When we consider regressions in

[7]Fisher (1959), Houweling, Mentink, and Vorst (2004), Longstaff, Mithal, and Neis (2004) and Perraudin and Taylor (2002) document the existence of an illiquidity component in bond yield spreads. In addition, Elton, Gruber, Agrawal, and Mann (2001) suggest that both the differential taxation of corporate and government bonds as well as compensation for systematic risk will impact bond spreads over and above the size of expected losses given default.

levels, explanatory power is quite high with R-squares ranging from 50% to 75%. Thus variables drawn from economic theory are clearly important in explaining the pricing of this particular type of credit-sensitive instrument. This finding is reinforced by an analysis of the regression residuals, which shows that the evidence for a remaining common component is weaker than in the work of CGM on corporate bond data. We argue however that a comparison of our results with empirical results on corporate bond spreads should be interpreted cautiously. One reason is that the particular maturity structure of the CDS data is likely to influence our conclusions on the explanatory power of the results.

The paper proceeds as follows. In the next section, we lay out our analytical framework. In particular, we discuss the determinants of default swap premia suggested by existing theory and then present our regression equations. In section 3, we present and discuss our empirical results. Section 4 concludes.

2 Analytical Framework

2.1 The Theoretical Determinants of Credit Default Swap premia

There are two different approaches to modeling credit sensitive financial instruments. One approach is due to Merton (1974) and relies on a theoretical approach that explicitly relates the credit event to the value of the firm's assets. The firm is assumed to default on its obligations when the firm value falls below some threshold. These types of models are called structural models because the link with economic fundamentals is explicit. They can be used to price credit sensitive securities such as corporate bonds as well as credit default swaps. The second approach is more recent and finds its origins in the modeling of the risk free term structure. This approach is referred to as the reduced form approach because the relationship with underlying economic variables such as the firm value is not explicitly modeled.

This paper analyzes CDS premia from the perspective of structural form models. Following Merton's (1974) pathbreaking work, the basic structural model has been extended in different ways.[8] While these models typically focus on the importance of additional theoret-

[8] See Black and Cox (1976), Geske (1977), Fischer, Heinkel, and Zechner (1989), Kim, Ramaswamy, and Sundaresan (1993), Nielsen, Saá-Requejo, and Santa-Clara (1993), Leland (1994), Longstaff and Schwartz (1995), Anderson and Sundaresan (1996), Leland and Toft (1996), Mella-Barral and Perraudin (1997), Zhou

ical variables, or change the precise functional dependence of default on existing theoretical variables, they all have in common that default and therefore the value of the default sensitive security depends on a number of determinants that are central to the Merton (1974) approach. First, leverage is central to all these models: ceteris paribus, the more levered the firm, the higher the probability of default. Second, the volatility of the underlying assets is an essential determinant of the value of the default sensitive security because the latter is equivalent to a credit risk free security and a short put. Volatility influences the value of the put option. Third, the level of the riskless rate also impacts the value of the option. Although the correlation between the risk free rate and the bond spread is strictly not part of Merton's (1974) analysis which relies on a constant interest rate, the framework does predict a negative relationship between these two variables. The reason is that the risk free rate determines the risk adjusted drift of firm value and thus an increase in this variable will tend to decrease risk adjusted default probabilities and also spreads. The same result has been shown in models where the dynamics of the risk free rate have been modelled explicitly.[9]

Rather than carrying out a full structural estimation of any given model or set of such, we rely on what these models together suggest are the main determinants of credit risk. We use these variables in simple linear regressions of default swap premia on the suggested factors. Note that although structural models have almost exclusively been used to value corporate bonds, the implied relationship between the theoretical variables and default swap premia is the same. This can be understood by considering the similarity between the payoffs of the two types of financial instruments. Bonds pay regular coupons and principal cash flows until default occurs. At that time, the bond will be worth a fraction of its original principal amount.[10] The seller of default insurance through a CDS (analogous to the holder of the bond) receives regular payments (approximately the coupon rate on the bond minus the risk free rate) until default occurs, when he makes a payment equivalent to the loss in market value of the underlying bond - thus incurring the same loss as the holder of the bond.[11]

(1997), Leland (1998), Mella-Barral (1999), Duffie and Lando (2000), Collin-Dufresne and Goldstein (2001), François and Morellec (2004).

[9]See e.g. Longstaff and Schwartz (1995) and Collin-Dufresne and Goldstein (2001).

[10]Models differ in their exact technical treatment of this payment but this is without implications for the differences between bonds and default swaps.

[11]In practice, the settlement in the event of default may be made either in cash or in kind. If made in cash, a third party typically determines the post credit event market value of the reference obligation according

Thus in terms of the sequence of cash flows and the impact of default, bonds and CDSs are very similar and structural model variables will have the same impact on the values of both securities.[12]

In what follows, we will study the link between theoretically motivated determinants of default risk and market data on CDS premia using simple linear regression methods. In doing so, we closely parallel the approach taken by CT, CGM and CDMW using corporate bond data.

2.2 Regressions

According to theory, the premia on credit default swaps should be determined by the amount of leverage incurred by the underlying firm, the volatility of the underlying assets and the riskless spot rate. We denote the leverage of firm i at time t as $lev_{i,t}$ and the volatility as $vol_{i,t}$. We define the riskfree rate variable to be the 10-year yield, denoted as r_t^{10}. This choice is motivated as follows. Theoretical models tend to be based on the dynamics of the instantaneous risk free rate, which is unobservable. A number of empirical studies have demonstrated that this unobservable short rate can be thought of as being determined by a number of factors, one of which is the yield on long-maturity bonds. In the interest of parsimony in the empirical presentation of the results, we therefore focus exclusively on this one proxy for the riskless spot rate in our base case regression results. The robustness of our findings with respect to a different choice of factor or the inclusion of additional factors is discussed in detail in Section 3.3.

The regression suggested by theory consists therefore of regressing the CDS premium, denoted by $S_{i,t}$, on these three variables. We also add a constant to this regression which

to a predetermined formula and the payment made will be the original principal minus this value. If the settlement is in kind, the buyer of insurance will put the bond to the seller at par. In some cases, there may be a certain amount of flexibility for the buyer as to which bond can be delivered, much like for government bond futures contracts.

[12] In fact, in the absence of counterparty risk and market frictions, it can be shown that a CDS on a floating rate bond originally issued at par can be synthesized by an offsetting portfolio of this floater and an otherwise identical credit risk free floater. The net cash flows of this portfolio must equal those of the CDS in the absence of arbitrage. See Duffie and Singleton (2003) for a detailed discussion of this and more complex cases.

yields

$$S_{i,t} = \alpha_i^l + \beta_i^l lev_{i,t} + \beta_i^v vol_{i,t} + \beta_i^r r_t^{10} + \varepsilon_{i,t}. \tag{1}$$

We also regress the premium on each of these regressors separately to get a better idea of the explanatory power of each regressor

$$S_{i,t} = \alpha_i^l + \beta_i^l lev_{i,t} + \varepsilon_{i,t}. \tag{2}$$

$$S_{i,t} = \alpha_i^l + \beta_i^v vol_{i,t} + \varepsilon_{i,t}. \tag{3}$$

$$S_{i,t} = \alpha_i^l + \beta_i^r r_t^{10} + \varepsilon_{i,t}. \tag{4}$$

CT and CDMW use similar regressions to investigate the importance of these theoretical variables for the determination of credit spreads on corporate bonds. CGM run similar regressions, but focus on changes in credit spreads. A specification in differences can be motivated statistically, because first differencing is appropriate if the dependent variable and regressors are integrated. Due to the relatively short time period covered by our data, it is difficult to verify which statistical assumption is appropriate. We therefore proceed to test the theory both in differences and levels. It must be noted that differences are harder to explain than levels and a regression in differences therefore should provide a more stringent test of the theory. We therefore also estimate regressions (1)-(4) in differences.

$$\Delta S_{i,t} = \alpha_i^d + \beta_i^l \Delta lev_{i,t} + \beta_i^v \Delta vol_{i,t} + \beta_i^r \Delta r_t^{10} + \varepsilon_{i,t}. \tag{5}$$

$$\Delta S_{i,t} = \alpha_i^d + \beta_i^l \Delta lev_{i,t} + \varepsilon_{i,t}. \tag{6}$$

$$\Delta S_{i,t} = \alpha_i^d + \beta_i^v \Delta vol_{i,t} + \varepsilon_{i,t}. \tag{7}$$

$$\Delta S_{i,t} = \alpha_i^d + \beta_i^r \Delta r_t^{10} + \varepsilon_{i,t}. \tag{8}$$

3 Empirical Analysis

3.1 Data

To investigate the regressions suggested by theory, we require data on credit default swap premia, firm leverage, volatility and riskless yields. We obtain these data from the following sources:

Credit Default Swap premia: We use quotes from the CreditTrade Market Prices database for 1999-2002 corresponding to credit default swaps on senior debt. The CDS market has experienced considerable growth over this period. Figure 1 depicts the evolution of the number of daily available quotes.

Only the contracts on companies for which we have data in CRSP and COMPUSTAT are used in our study. The North America Industry Classification System (NAICS) code was obtained for each company from FISD and WRDS. Using the NAICS code, utilities and financial companies were excluded. Since there are very few quotes on junior debt, these quotes are excluded. The amount of quotes satisfying the above criteria is 53,625. At least 92% of these quotes are firm, in that they represent a committment to trade a given notional principal ranging from $2 million to $10 million.[13] Figure 2 depicts the number of quotes as a function of the tenor. The market is clearly concentrated on maturities around 5 years. We therefore only retain 48,626 quotes that have tenors between 4.5 and 5.5 years. This sample represents 90.7% of all quotes.

Even though the CDS market is a worldwide market, the majority of the quotes fall within New York trading hours. This finding is to a large extent due to our selection criteria, because CRSP and COMPUSTAT mainly contain data on US companies. From the 48,626 quotes, we selected, for each day and reference entity, the quote closest to 4PM NY time. More precisely, we filter the quotes according to the following criteria:

- Either the time stamp is after 3PM
- or the time stamp is between 12 noon and 3PM and the time stamp on the previous available quote is more than two trading days old

[13] The remaining 8% of the quotes are recorded with a zero notional amount. However, according to personal communication with CreditTrade staff, a zero notional simply indicates that it was not recorded.

- or the time stamp is between 9AM and 12 noon and the time stamp on the previous available quote is more than three trading days old
- or the time stamp is between 6AM and 9AM and the time stamp on the previous available quote is more than four trading days old
- or the time stamp is between 3AM and 6AM and the time stamp on the previous available quote is more than five trading days old.

This rule is motivated by consideration for the difference regressions. To compute the differences in the premia, we ideally want quotes at the exact same time of the day. This is not possible and because of sample size considerations, it is also not possible to limit ourselves to time stamps after 3PM. By including quotes with time stamps further removed from 4PM, the potential for biases in the computed premium differences increases. However, by only selecting quotes farther removed from 4PM if the previous quote is further removed in time, we ensure that the potential bias from time stamps at different parts of the day is reduced.

Bid and offer quotes are treated separately.[14] As a final filter, we only retain firms with at least 25 quotes or changes in quotes, depending on the regression specification. It should be noted that the number of observations in any given regression will depend on whether it is run on levels or differences and on whether bids or offers are used. This leaves us with 4813 bid and 5436 offer quotes over the whole sample period, with slightly fewer observations for regressions in differences. The Appendix lists the companies that are included in the sample for the different regressions.

The data for the theoretical determinants of the CDS premia (the explanatory variables in the regressions) are constructed as follows:

Leverage: The leverage ratio is defined as

$$\frac{\text{Book Value of Debt} + \text{Book Value of Preferred Equity}}{\text{Market Value of Equity} + \text{Book Value of Debt} + \text{Book Value of Preferred Equity}} \tag{9}$$

[14] We obtain very similar results when including bids and offers in the same regression. However, in that case, the bid-ask bounce affects the explanatory power of the regression. When including a dummy to correct for this, the R-squares increase as the variation in the bid-ask spread is captured by the dummy. We therefore prefer to report results on separate regressions for bids and offers.

The Market Value of Equity was obtained from CRSP, and the Book Value of Debt and the Book Value of Preferred Equity from COMPUSTAT. Since book values are only available at the quarterly level, we linearly interpolate in order to obtain daily figures.

Volatility: A time series of equity volatility was computed for each company using an exponentially weighted moving average model on daily returns obtained from CRSP.[15] In the empirical literature on the determinants of corporate bond spreads, our approach is closest to that of CT, who construct historical volatility based on 180 days of returns in their base case regressions. CGM use the VIX data, which represents option-implied volatility based on S&P 100 index options. CDMW use both volatility implied by individual equity options as well as historical volatility.

Treasury Bond Yields: Daily data on 10-year Treasury bond yields were collected from DataStream. We use the appropriate constant maturity index constructed by the US Treasury based on the most actively traded issues in that maturity segment.

Table 1 and Figure 3 provide descriptive statistics and visual summaries of the CDS premia and the explanatory variables used in the main regressions. The CDS premium is 180 basis points on average with a large standard deviation. The explanatory variables seem to be less variable than the CDS premium and especially the 10-year yield is tightly centered around the mean. From Figure 3 it would seem that the high variability of the CDS premium is partly due to the fact that the premium has been increasing over time, regardless of the rating of the reference obligation, and that the premium differs considerably across reference obligations with different ratings. Figure 3 also clearly indicates that the number of available datapoints is very different for different reference obligations.

Because the data set has a cross-sectional as well as a time-series dimension, several aspects of the relationship between the theoretical variables and the credit spreads can in principle be investigated. Cross-sectional correlations indicate how credit spreads differ between companies because of differences in leverage and volatility. Time-series correlations indicate how credit spreads change for a given company as the company's leverage ratio and equity volatility change. Table 1 presents some initial insight into these correlations and the differences between the cross-sectional and time-series patterns. Time-series as well as

[15] For each reference entity, volatility h_t was generated according to $h_t = r_t^2(1-\lambda) + h_{t-1}\lambda$, with r_t denoting daily returns. In order to obtain a more precise estimate of λ, we constrain this parameter to be the same across firms in the estimation.

cross-sectional correlations between the CDS premia and the theoretical variables have the expected sign, and interestingly for both volatility and leverage the cross-sectional correlation is not very different from the time-series correlation. Figures 4 through 7 provide additional insight into this issue. Figures 4 and 5 graphically illustrate the time-series relationship with the CDS premia for leverage and volatility respectively by averaging the variables across firms at a point in time. Because our data are unevenly spaced, we use weekly averages. The figures clearly suggest a positive time-series relationship between either theoretical variable and the CDS premium. Figures 6 and 7 graphically illustrate the cross-sectional relationship between the variables and the CDS premia by averaging the data across time for a given firm. While the figures suggest a positive relationship for volatility as well as leverage, they clearly confirm the result in Table 1 that the correlation is higher for volatility.

3.2 Regression Results

Because the regressions (1)-(8) have a cross-sectional as well as a time-series dimension, they can be implemented in different ways. We first follow CGM and present results on average regression coefficients obtained by running a series of time-series regressions for every different company, emphasizing time-series correlations between CDS premia and theoretical variables. From a managerial perspective, these regressions are of most interest because they indicate how credit spreads change for a given company as the company's leverage ratio and equity volatility change. Subsequently, for the levels regressions (1)-(4) we also present results obtained using a number of different panel data techniques. Regarding the implementation of the regressions, note that the constant in the difference regressions is obviously different (at a theoretical level) from the constant in the levels regressions, which is why it is indexed with a superscript l or d, respectively. The constant is also indexed by a subscript i, because in the implementation using time-series regression it is different for every company. In the panel data implementation, this is the case when estimating fixed-effects but not for the OLS panel data regression which constrains the constant to be the same for all companies.

Table 2 presents the results of the levels regression (1) and the difference regression (5). For both regressions, we report results obtained with bid quotes as well as results obtained with offer quotes. In each case, we report results obtained using data on all companies, and we also report results for a sample of companies with below median rating and another sample of companies with above median rating. The number of companies included in each analysis is listed in the third row from the bottom. The next to last row indicates the average number of observations included in the time-series regressions, and the last row indicates on average how much time elapses between different quotes for the same underlying. For each case, the top four rows list the average regression coefficients obtained from the time-series regressions. Rows 5 through 8 present the t-statistics, computed in the same way as in Collin-Dufresne, Goldstein, and Martin (2001). They therefore represent cross-sectional variation in the time series regression coefficient estimates.

A number of important conclusions obtain. First, the estimated sign for the coefficient on leverage is always positive, as expected a priori. Second, the estimated sign for the coefficient on volatility is also always positive, as expected. Third, the coefficient on the 10-year yield also conforms to theoretical expectations because it is estimated with a negative sign. What is even more encouraging is that the t-statistics almost uniformly indicate statistical significance at conventional significance levels. Interestingly, the few exceptions occur for the levels regressions, not for the (more challenging) difference regressions.

The point estimates for the coefficients are remarkably similar across the levels and difference regressions, least so for the coefficients on the 10-year yield. Not surprisingly, there are some differences in the point estimates across ratings. For lower rated firms, the point estimates for leverage and volatility are bigger than for higher rated firms. These effects are perfectly intuitive and consistent with the predictions of any structural credit risk model. We also find that CDS premia for lower rated firms are more sensitive to interest rates. Again, this is consistent with the theory. It is also consistent with the empirical findings of Duffee (1998) on corporate bond yield spreads.[16]

A final statistic of interest is the adjusted R^2. First and foremost, the explanatory power of the levels regressions is of course much higher than that of the difference regressions. For the levels regressions, the theoretical variables explain approximately 60% of the variation in the premium. For the difference regressions, the theoretical variables explain approximately 23%. The R-squares for the lower ratings are always a bit higher than those for the higher ratings, as expected. It may also be of interest that in the level regressions the R-squares

[16] In a structural model, the risk adjusted probability of default is decreasing in the risk free interest rate. Intuitively, a higher risk free rate entails a higher drift rate for the firm's asset value and allows it to grow its way away from financial distress. See also Longstaff and Schwartz (1995).

for the bid quotes are a bit higher than the R-squares for the offer quotes, even though this pattern does not show up in the difference regressions.

While the effects of a change in the yield curve somewhat depend on whether one estimates in levels or differences, the results for volatility and leverage are robust across specifications. This renders the economic interpretation of the point estimates of significant interest. Using the estimation results for all companies, an 1% increase in (annualized) equity volatility raises the CDS premium on average by approximately 0.8-1.5 basis point. For companies with lower ratings, the effect is estimated to be between 1.1 and 2.3 basis points. The leverage effect is also stronger for lowly rated companies: a 1% change in the leverage ratio increases their CDS premium by approximately 6-10 basis points, whereas this effect is between 4.8 and 7.3 basis points when considering all companies.

Tables 3, 4 and 5 further explore these results. Table 3 presents results for regressions (2) and (6), Table 4 for regressions (3) and (7), and Table 5 for regressions (4) and (8). The tables are structured in the same way as Table 2. It must be noted that, in a sense, the point estimates in these tables are of somewhat less interest than those in Table 2, because the regression in Table 2 is the one suggested by the theory. It is therefore entirely possible that in the univariate regressions in Tables 3 through 5, coefficients are biased because of an omitted variable argument.

Interestingly however, the signs of the point estimates are the same as in Table 2 and the t-statistics for the time-varying regressors are significant at conventional significance levels. Table 3 indicates that when leverage is the only explanatory variable, its economic effect is always estimated to be larger than in Table 2, and the same is true for volatility in Table 4, but the effects are roughly of the same order of magnitude. A comparison of the R-squares in Tables 3-5 with those in Table 2 indicates to what extent each of the theoretical variables contributes to the explanatory power of the regression. It can be seen that each of the three variables has some explanatory power, even though the leverage variable clearly dominates the other two regressors. The leverage variable alone explains between 37.1% and 45.7% of the variation in CDS premia in the levels regressions, but only about 13% on average in the difference regressions. Volatility explains between 23.9% and 29.7% in the levels regressions, but only between 6.9% and 14.4% in the difference regressions. Interestingly, the 10-year yield variable has a higher R-square than the volatility variable in the levels regression, but its explanatory power in the difference regressions is decidedly modest.

Note that the negative correlation between CDS premia and the risk free rate discussed above has also been documented for bond yield spreads by Longstaff and Schwartz (1995) and Duffee (1998). Presently, no consensus prevails as to the economic reasoning behind this stylized fact. Duffie and Singleton (2003) state that one possible explanation for the negative correlation is the existence of stale corporate bond prices. The spreads are measured by taking the difference between the corporate and the Treasury yield curves; therefore, an increase in Treasury yields might be associated with a decrease in spreads until the recorded corporate bond price accounts for the change. Our results rule out the latter explanation because default swap premia are not given by the difference of two yields as bond spreads are. However, our results are consistent with the implication of structural models that an increase in the risk free rate will decrease risk-adjusted default probabilities.[17]

In summary, we conclude that there are some interesting differences between the levels and difference regressions in Tables 3-5.

3.3 Robustness Analysis

This Section further investigates the robustness of the regression results presented in Section 3.2. In a first step, we estimate the regression proposed by CGM. Their base case regression includes the explanatory variables $lev_{i,t}$, $vol_{i,t}$ and r_t^{10} included in (1) but adds a number of other explanatory variables including

Treasury Bond Yields: We collected daily series of 2-year and 10-year bond yields from DataStream.

The Slope of the Yield Curve: Defined as the difference between the 10-year Treasury bond yield used in regression (2) and 2-year Treasury bond yields also obtained from DataStream. We use the 2-year Treasury bond yield as the level of the yield curve in order to make the interpretation of the slope more straightforward.

The square of the 2-year yield.

The return on the S&P 500: Daily data on the return on the S&P 500 was obtained from DataStream.

The slope of the smirk: We estimate the slope of the smirk on equity options using out-of-the-money S&P 500 American futures put options from the CME Futures and Options

[17] See Longstaff and Schwartz (1995) for a discussion.

Database. A number of choices have to be made as regards these calculations. First, implied volatilities are computed using the American options analytical approximation technique proposed by Whaley (1986). Second, we cannot simply compute the smirk using one particular maturity because the same maturity is not available on every trading day. To take into account the dependence of the smirk on maturity, we define moneyness as $ln(K/F)/sqrt(T)$, were K is the strike price, F is the futures price, and T is the time to expiration. Standardizing moneyness by $sqrt(T)$ makes the slope of the smirk (on a given trading day) remarkably similar across expirations. Third, we estimate a linear relation between moneyness and implied volatility.[18] Robustness tests demonstrate that adding a quadratic term does not change the results. Fourth, we arbitrarily choose 45 days as a benchmark maturity. The slope of the 45-day smirk is then obtained from linearly interpolating the coefficients corresponding to the nearest available expirations.

The motivation for including these variables is as follows. The interest rate variable directly modeled by most of the theory is the instantaneous spot rate. It has been shown empirically that the instantaneous rate can be explained by a number of term structure variables. The yield on long maturities used in regression (1) is one of these variables. Alternatively, one can use the yield on short maturity bonds or the difference in yield between short and long maturities, which is what is proposed here. The square of the 2-year yield is a convenient attempt to exploit nonlinearities in the relationship between term structure variables and credit default swap premia. CGM (2001) use the return on the S&P 500 to proxy for the overall state of the economy and the slope of the smirk to proxy for jumps in firm value. It is clear that some of these variables are more loosely related to theory compared to the regressors in (1). For additional motivation see CGM. Including these explanatory variables leads to the levels regression

$$S_{i,t} = \alpha_i^l + \beta_i^l lev_{i,t} + \beta_i^v vol_{i,t} + \beta_i^r r_t^2 + \beta_i^{r2} (r_t^2)^2 + \beta_i^{r3} tsslop_t + \beta_i^{sp} S\&P_t + \beta_i^{sm} smslop_t + \varepsilon_{i,t}. \quad (10)$$

and the difference regression

$$\Delta S_{i,t} = \alpha_i^d + \beta_i^l \Delta lev_{i,t} + \beta_i^v \Delta vol_{i,t} + \beta_i^r \Delta r_t^2 + \beta_i^{r2} (\Delta r_t^2)^2 + \beta_i^{r3} \Delta tsslop_t + \beta_i^{sp} \Delta S\&P_t + \beta_i^{sm} \Delta smslop_t + \varepsilon_{i,t} \quad (11)$$

[18] To circumvent the noise in very deep out-of-the-money options, we ignore options whose moneyness was lower than the median across time of the lowest moneyness of each trading day.

Table 6, which presents the results of these regressions, has the same format as Table 2. The t-statistics were computed in the same fashion. One objective of this table is to verify by means of the R-squares if the addition of these variables increases the explanatory power of the theory. For the difference regressions, the extra variables increase the R-square by roughly 7.5%, whereas for the levels regressions the increase in the R-square is approximately 14%. Interestingly, the increase in R-square is larger for the regressions that use offer quotes. The term structure variables are often insignificantly estimated, perhaps suggesting some multicollinearity between them, or high correlation with another explanatory variable. The return on the S&P 500 has a significantly estimated negative impact on the CDS premium, indicating that in times with high returns (good times), the premium narrows. This finding is consistent with the findings in CGM for spreads on corporate bonds. The slope of the smirk seems to have a minor impact on the CDS premium. Finally and perhaps most importantly, the point estimates for leverage and volatility are very similar to the ones in Table 2. We therefore conclude that the magnitude of the effects discussed before is robust to the inclusion of a number of other variables. This is remarkable if one considers that the R-square increases considerably, and that in the levels regression we have a specification in Table 6 that explains a large part of the variation in CDS premia. These results therefore inspire confidence in our estimates.

It could be argued that the t-statistics in Tables 2 through 6 are hard to interpret because they are computed based on the variation in regression coefficients for time-series regressions. An alternative approach is to treat the empirical problem as a full-fledged panel data problem. Tables 7 and 8 present the results of this procedure for the levels regressions. We do not report panel estimation for the difference regressions because we need a number of additional assumptions regarding cross-sectional correlation patterns and autocorrelations to compute standard errors, and the levels regressions are sufficient to make the point.

For all three panels in Tables 7 and 8, columns 1-4 report results for estimation of regressions (1), (2), (3) and (4). Table 7 reports results for offer quotes and Table 8 for bid quotes. Panel A reports results for the basic OLS panel regression, Panel B allows for reference entity fixed effects and Panel C includes quarter dummies. In Panel A, each observation is treated independently and the regression constant is assumed to be the same across companies.[19]

[19] An advantage of the panel data approach is that it can be applied without imposing a minimum number

The point estimates in Panel A again have the signs predicted by theory, although their magnitudes differ from the firm by firm time-series regressions in Tables 2-6. The coefficient for leverage tends to be smaller while equity volatility enters with a larger coefficient. When, as in Panel B, fixed effects for the reference entities are included, the parameter estimates fall back in line with what was found in Tables 2-6, while the R-squares increase substantially. This clearly indicates that there is a large amount of cross-sectional variation that cannot be captured by the theoretical variables. The main effect of including quarter dummies (Panel C) is a slight increase in the R-square of the regression relative to the base case in Panel A. This can be interpreted as suggesting that the theoretical variables explain most of the time-series variation in the data, but the results of this regression may be hard to interpret. The results will be affected by inserting more time dummies into the equation, and the choice of quarterly dummies is ad hoc. Note however that because we have daily data and an unbalanced panel, there is no natural choice for the frequency of the time dummies: quarterly dummies are as good a choice as any. The t-statistics are much higher in Tables 7 and 8, which is not necessarily surprising because the t-statistics in Tables 2-6 are essentially computed on the variation in the regression coefficients and therefore hard to relate to the more conventional t-statistics in Tables 7 and 8.[20]

Despite some of the problems with the interpretation of the time dummies, the relative increases in explanatory power resulting from including fixed effects and time dummies respectively suggests that the variables determined by theory may have more explanatory power in a time series than a cross-sectional sense. In this respect, it is interesting to note that the ranking of the R-squares for the univariate regressions on leverage and volatility (in Tables 3, 4, 7 and 8) differs depending on whether the data is treated as a collection of time series or as a panel. In the time series case, the R-squares are higher when leverage is used as regressor compared to equity volatility. In Panels A of Tables 7 and 8, the equity

---

of observations per reference entity. Although we report results for data that conform to that used in the time series regressions of Tables 2-6, results are very similar when we remove the requirement that firms have a minimum of 25 quotes. We then have a total of 13033 bids and offers, about 27% more than with the restriction.

[20] We also run separate regressions for 1999, 2000, 2001 and 2002 in order to investigate the robustness of the results over time. Estimated coefficients are robust over time for volatility and leverage. However, for 1999, the economic significance of these variables is smaller. In addition, an interesting result is that the R-square increases noticeably over time.

volatilities appear to be more successful in explaining the variation in CDS premia. This is consistent with leverage having more explanatory power in the time series, whereas volatility is *relatively* speaking better at explaining the cross section. Returning to Table 1, we can see that the reported cross-sectional and time series correlations are consistent with this observation. For leverage, the cross-sectional correlation is lower than the time-series correlation, while it is the opposite for volatility. One possible explanation is that theoretically, leverage does not provide sufficient information about the likelihood of financial distress since it does not convey information about business risk. Equity volatility, on the other hand, provides information about both asset risk and leverage, and can thus be better used to discriminate between the credit risk of different firms.

Finally, note that the cross-sectional and time-series correlations in Table 1 also help to explain the differences in the point estimates between Panel A of Tables 7 and 8 on the one hand and Panel B (as well as Tables 2-4) on the other hand. The fixed effects regressions in Panel B capture the time-series correlation. Because the results in Panel A capture a mixture of time-series and cross-sectional correlation, the point estimate for leverage goes down and that for volatility goes up, consistent with the relative strength of the effects documented in Table 1. It is interesting to note that the small differences between the time-series and cross-sectional correlations documented in Table 1 leads to relatively large changes in point estimates.

3.4 Discussion

It is interesting to compare these results with the results obtained for spreads on corporate bonds by CT, CGM and CDMW. The most important observation is that our results confirm the results in these papers that the theoretical determinants of credit risk are empirically relevant and estimated with the sign predicted by theory. With respect to the explanatory power of these theoretical variables, a comparison is unfortunately less straightforward. CGM use a market-wide measure of volatility. They estimate difference regressions and their base-case regressions are the ones in Table 6. The R-squares in CGM are considerably lower. They also obtain much lower R-squares than we do when studying the effects of leverage in isolation. Our point estimates for the effects of leverage and volatility are larger than theirs, but it must of course be noted that our measure of volatility is very different. CT investigate level regressions and focus mainly on the effect of volatility. They also use a historical measure of volatility and because they use panel regressions their results are most closely related to those of Tables 7 and 8. In general they obtain higher R-squares than we do, but this finding must be interpreted with caution because they include a number of control variables which explain approximately 25% of all variation. The estimate of a 1% change in annualized volatility in CT is 14 basis points, considerably higher than our estimate.

Some of the empirical results in CDMW are closely related to the ones in this paper because they investigate the explanatory power of volatility in the absence of other explanatory variables. However, they do not consider the impact of leverage. CDMW use panel regressions and the R-squares and point estimates in their base-case regressions ought to be compared to the ones in Tables 7 and 8. It is noteworthy that their point estimates for the firm implied volatility are very similar to the ones we obtain using historical volatility. This is likely due to the fact that we compute volatility as an exponentially weighted moving average, which like implied volatility is more variable than a 180 day historical average.

In summary, the explanatory power of the theoretical variables in our analysis differs from the results in the literature on corporate bond spreads, which itself contains some divergent results. It must be noted that it may be problematic to try to relate the explanatory power of regressions for corporate bond spreads to those for CDS premia. The reason is that the explanatory power of the regressions depends on maturity (see CT, CGM and CDMW). Because the maturity of the Credit Default Swaps in our sample (roughly five years) may be very different from the average maturity for corporate bonds, this may compromise a comparison of R-squares between the two markets.

3.5 Analyzing the Regression Residuals

One robust conclusion from Tables 2-8 is that the theoretical determinants of CDS premia are estimated statistically significantly with signs that confirm our intuition and that the magnitude of these effects is also intuitively plausible. However, it is difficult to determine how successful theory is in explaining the variation in CDS premia. The R-squares of the explanatory regressions vary considerably dependent on whether one analyzes levels or differences, and on whether one uses panel data or time series techniques. Moreover, we do not necessarily have good benchmarks for the R-squares, because comparisons with empirical results for the corporate bond market are subject to problems.

We therefore attempt to provide more intuition for the explanatory power of the theoretical determinants of CDS premia. To understand the structure of the remaining variation in the data after controlling for the theoretical determinants of CDS premia, we analyze the regression residuals from the levels regression (1) and the difference regression (5) using principal components analysis (PCA). By analyzing the correlation matrix of the errors of the time-series regressions, we investigate if there exists an unidentified common factor that explains a significant portion of the variation of the errors. The structure of the data somewhat complicates the analysis, and we performed a number of different analyses in order to investigate the robustness of our conclusions. There are two types of complications in the data. First, the data are non-synchronous. Second, the number of observations differs considerably by company. The first complication causes some difficulties at a technical level. The second complication forces us to make some choices regarding the use of the data.

We first report on an analysis of the levels regression (1), using the correlation matrix of the regression errors for the 15 companies with the highest number of observations. We limit ourselves to a small number of companies to obtain results that are based on as much time-series information as possible. We also analyze the correlation matrix of the CDS premia $S_{i,t}$. For premia and errors from the levels regressions, a simple approach to the non-synchronicity problem is available. We artificially construct observations every 7 calendar days, by linearly interpolating from the closest (in time) two observations. This results in a balanced panel of errors. Panel A of Table 9 shows that for the bid quote levels, the first principal component is fairly important, explaining 58.7% of the variation. The first eigenvector has mostly positive elements of similar magnitude, with a few exceptions. The first principal component of the errors has more diverse weights, and it explains only 32.5% of the variation of the errors. The results for offer quotes in Panel B support those from Panel A. The first principal component for the errors explains only 31.0% or the error variation. The difference between the explanatory power of the first principal component of the premium difference and that of the errors is approximately 25%, similar to the difference in Panel A. A comparison between these R-squares suggests that a substantial part of the common variation of the premia is explained by the regressors.

Table 10 repeats the analysis of Table 9, using the 15 companies with the highest number of observations, but uses the errors of the time-series regressions in differences (5). For differences, a simple interpolation does not work because there is more than one time index. Instead, each element of the correlation matrix has to be estimated individually. We do so by using the procedure of de Jong and Nijman (1997).[21] Because the estimated correlation matrix is not generally positive semidefinite, we compute the positive semidefinite matrix closest to the estimated correlation matrix according to the Frobenius-norm using a numerical algorithm due to Sharapov (1997) and also used by Ledoit, Santa-Clara, and Wolf (2003).

Panel A of Table 10 shows that for the differences in bid quotes the first principal component is fairly important, explaining 50.2% of the variation, with a first eigenvector that has only positive elements. In contrast, the first principal component of the errors has positive and negative elements, and it explains only 24.5% of the variation of the errors. The results for offer quotes in Panel B are a bit weaker but support those from Panel A. In this case the first principal component of the errors contains only one negative element, but the weights of the first principal component of the differences in offer quotes are remarkably more homogeneous. Most importantly, the first principal component for the errors explains only 30.8% or the error variation.

Our third PCA is closer in spirit to the one in CGM, although it is slightly different because of data constraints. CGM perform a PCA by distributing the errors of all the companies in the sample in bins according to the maturity of the bonds and the leverage of the issuing companies. With a balanced panel, it is straightforward to do this analysis for differences. In our case, we do not observe the premia at fixed intervals. As a result, changes in premia and the corresponding errors carry a double time index, and it is not feasible to assign them to bins. We therefore limit ourselves to a PCA using bins for the levels regressions (1).

CGM construct fifteen bins by classifying the companies in 5 leverage groups and the bonds in three maturity ranges. However, because all CDSs in our sample have (roughly) a 5-year maturity, it is not feasible to use maturity as a classification variable. Also, we have only one kind of CDS per company, and not a collection of bonds. Therefore we construct our bins using only the leverage dimension, so that we have 5 bins delimited by the quintiles

[21] Martens (2003) reviews and compares different methods for computing covariance matrices for nonsynchronous data. His simulations show that the de Jong and Nijman (1997) method is the most reliable in the absence of a bid-offer spread. Given that we work with either bids or offers, we choose this method.

of the distribution of leverage of the different companies. The time interval defining the bins is 15 days. Table 11 reports on this analysis. The first principal component for the bid (offer) errors explains only 35.6% (36.4%) of the variation of the bins, compared to 68.6% (66.1%) for the bid (offer) quotes.

Overall, the three tables allow for a remarkably robust conclusion. The PCAs for the levels and differences suggest that the theoretical determinants of default swap premia do explain a significant part of the common variation. Regarding the percentage of the variance explained by the first principal component in the error analysis, it varies dependent on whether one uses bins and whether one uses differences or levels, but it varies between 20% and 36%. A high percentage in this case would indicate that there is a lot of common variation left which cannot be explained by on of the theoretical variables. However, we find it difficult to draw strong conclusions from this range of numbers as to the validity of the theoretical variables, because it is not clear what the benchmark is. Compared with the findings in CGM, the percentage variation explained by the first principal component in the errors is certainly low. It must also be taken into account that the largest eigenvalues are in general severely biased upward, as observed by Ledoit and Wolf (2004).

To further understand the nature of the residuals, we also ran regressions (1) and (5) with a CDS market index included. One would expect such an index to have substantial explanatory power for residual CDS premia if the variables suggested by theory are inadequate. Unfortunately no index is available for the CDS market over our entire sample. We use the TRACERS index, which is available from September 2001 to the end of our sample and we repeat our estimation exercise with the CDS data available for this period (not reported).[22] It must be noted that although this covers less than half of the time period of our CDS sample, it covers the majority of the datapoints because the number of quotes increases through time. Interestingly, we find that including the market index does not noticeably affect the explanatory power of the regression. We therefore conclude that these results confirm those from Tables 9-11: the theoretical variables perform adequately in explaining CDS spreads.

[22] Morgan Stanley's TRACERS index is a synthetic index of US investment grade credit based on a selection of the most liquid reference entities.

4 Conclusion

Using a new dataset of bid and offer quotes for credit default swaps, we investigate the relationship between theoretical determinants of default risk and actual market premia. These determinants are firm leverage, volatility and the riskless interest rate. We find that these variables are statistically significantly estimated and that their effect is economically important as well as intuitively plausible. Moreover, the estimates of the economic effects of leverage and volatility are very similar regardless of whether one estimates on levels or differences and regardless of the econometric methodology. A 1% increase in annualized equity volatility raises the CDS premium by 1 to 2 basis points. A 1% change in the leverage ratio raises the CDS premium by approximately 5 to 10 basis points. These effects are not out of line with some of the estimates available in the literature on corporate bond spreads, even though Campbell and Taksler (2003) estimate a stronger effect of a change in volatility.

While these estimated effects are very robust and intuitively plausible, it is difficult to determine how successful the theory is in explaining the variation in the sample of CDS premia. The explanatory power of the theoretical variables depends on the econometric method and on whether one uses levels of differences. Using time series regressions the R-square for changes in default swap premia is approximately 23%, and the explanatory power for the levels of the premia is approximately 60%. The R-square for levels regressions goes up to more than 70% if we add in other explanatory variables as in Collin-Dufresne and Goldstein (2001). For a number of reasons it is difficult to relate these numbers to the available literature on other securities such as corporate bonds. However, our analysis of the residuals, coupled with the high R-squares for most of the levels regressions, leads us to cautiously conclude that the theory is successful in explaining the variation in CDS premia.

These results suggest a number of interesting questions. First, given that the variables critical for structural models of credit risk seem to be important for explaining CDS premia, how successful are structural models in explaining the data? One can think of the linear regressions in this paper as a first-order approximation to any structural model, suggesting that structural models may work well, but CT find that this logic does not extend to the Merton model when explaining corporate bond spreads. Second, an analysis of the effects of volatility based on individual equity options as in CDMW may prove worthwhile. Third, given that some of the estimated effects are very similar to those estimated in the corporate bond market, a further exploration of the interactions between the corporate bond market and the CDS market may prove worthwhile. Houweling and Vorst (2001) and Longstaff, Mithal, and Neis (2004) document some of these interactions using a reduced-form approach. It may prove worthwhile to explore the interactions between these markets by focusing on structural variables.

Appendix: Companies Used in Differences and Levels Regressions

Issuer Names	Differences		Levels	
	Bid	Offer	Bid	Offer
ABITIBI-CONSOLIDATED INC	1	0	1	0
ALBERTSONS INC	1	1	1	1
ALCOA INC	1	1	1	1
AMR CORP	0	1	1	1
AOL TIME WARNER INC	1	1	1	1
ARROW ELECTRONICS INC	1	1	1	1
AT&T WIRELESS SERVICES INC	1	1	1	1
AUTOZONE INC	0	1	0	1
BELLSOUTH CORPORATION	1	1	1	1
BLACK AND DECKER CORP	1	1	1	1
BOEING CO	1	1	1	1
BORGWARNER INC	0	1	0	1
BOSTON SCIENTIFIC CORP	0	1	0	1
BURLINGTON NORTHERN SANTA FE CORP	1	1	1	1
CAMPBELL SOUP CO	1	1	1	1
CARNIVAL CORP	1	1	1	1
CATERPILLAR INC	1	1	1	1
CENDANT CORP	1	1	1	1
CENTEX CORP	0	1	0	1
CENTURYTEL INC	0	1	0	1
CITIZENS COMMUNICATIONS CO.	1	1	1	1
CLEAR CHANNEL COMMUNICATIONS INC	1	1	1	1
COCA-COLA ENTERPRISES INC	1	1	1	1
COMPAQ COMPUTER CORP	1	1	1	1
COMPUTER ASSOCIATES INTERNATIONAL INC	0	1	0	1
CONAGRA FOODS INC	1	0	1	0
COX COMMUNICATIONS INC	1	1	1	1
CSX CORP	1	1	1	1
CVS CORP	1	1	1	1
DANA CORP	1	1	1	1
DEERE AND CO	1	1	1	1
DELPHI AUTOMOTIVE SYSTEMS CORP	1	1	1	1
DELTA AIRLINES INC	0	1	1	1
DILLARDS INC	1	0	1	0
DOW CHEMICAL CO, THE	1	1	1	1
DUPONT DE NEMOURS CO	0	1	0	1
EASTMAN KODAK CO	1	1	1	1
EL PASO CORP	1	1	1	1
ELECTRONIC DATA SYSTEMS CORP	1	1	1	1
ENRON CORP	1	1	1	1
FEDERAL EXPRESS CORP	0	1	0	1
FEDERATED DEPARTMENT STORES INC	1	1	1	1
GAP INC, THE	1	1	1	1
GENERAL MOTORS CORP	1	0	1	0
GEORGIA-PACIFIC CORP	1	1	1	1
GOODRICH CORP	1	1	1	1
GOODYEAR TIRE AND RUBBER CO, THE	1	1	1	1
HEWLETT-PACKARD CO	1	1	1	1
HILTON HOTELS CORP	1	1	1	1
HJ HEINZ CO	0	0	0	1
INGERSOLL-RAND CO	0	1	0	1
INTERNATIONAL BUSINESS MACHINES CORP	1	1	1	1
INTERNATIONAL PAPER CO	1	1	1	1
INTERPUBLIC GROUP COS. INC	1	1	1	1
JC PENNEY CO INC	1	1	1	1
KROGER	1	0	1	0
LENNAR CORP	0	1	0	1
LIMITED BRANDS	0	0	0	1
LOCKHEED MARTIN CORP	1	1	1	1
LUCENT TECHNOLOGIES INC	1	1	1	1
MASCO CORP	1	1	1	1
MATTEL INC	0	1	0	1
MAY DEPARTMENT STORES CO	1	1	1	1
MAYTAG CORP	1	1	1	1
MCDONALDS CORP	1	1	1	1
MCKESSON CORP	0	0	0	1
MGM MIRAGE INC	1	1	1	1
MOTOROLA INC	1	1	1	1
NEWELL RUBBERMAID INC	1	1	1	1

Appendix: Companies Used in Differences and Levels Regressions

Issuer Names	Differences		Levels	
	Bid	Offer	Bid	Offer
NEXTEL COMMUNICATIONS INC	1	1	1	1
NORDSTROM INC	1	1	1	1
NORFOLK SOUTHERN CORP	1	1	1	1
NORTHROP GRUMMAN CORP	0	1	0	1
OMNICOM GROUP	1	1	1	1
PARK PLACE ENTERTAINMENT CORP	1	1	1	1
PRIDE INTERNATIONAL INC	0	0	1	0
PROCTER AND GAMBLE CO, THE	0	1	0	1
ROYAL CARIBBEAN CRUISES LTD	1	0	1	1
RYDER SYSTEM INC	0	1	0	1
SBC COMMUNICATIONS INC	1	1	1	1
SEARS ROEBUCK AND CO	1	1	1	1
SOLECTRON CORP	1	1	1	1
SOUTHWEST AIRLINES CO	1	1	1	1
SPRINT CORP	1	1	1	1
SUN MICROSYSTEMS INC	1	1	1	1
TARGET CORP	1	1	1	1
TENET HEALTHCARE CORP	1	0	1	0
TJX COMPANIES INC	0	1	0	1
TOYS R US INC	1	1	1	1
TRIBUNE CO	0	1	0	1
TRW INC	1	1	1	1
TYCO INTERNATIONAL LTD	1	1	1	1
VIACOM INC	1	1	1	1
VISTEON CORP	1	1	1	1
WAL-MART STORES INC	1	1	1	1
WALT DISNEY CO, THE	1	1	1	1
WEYERHAEUSER CO	0	1	0	1
WHIRLPOOL CORP	1	1	1	1
WILLIAMS COMPANIES INC	1	1	1	1
WYETH (AMERICAN HOME PRODUCTS CORP)	1	1	1	1
XEROX CORP	1	1	1	1

References

Anderson, Ronald, and Suresh Sundaresan, 1996, Design and valuation of debt contracts, *Review of Financial Studies* 9, 37–68.

Bakshi, Gurdip, Dilip Madan, and Frank Zhang, 2001, Investigating the sources of default risk: Lessons from empirically implementing credit risk models, *Working paper, University of Maryland.*

BBA, 2003, *Credit Derivatives Report 2002/2003* (British Bankers' Association).

Berndt, Antje, Rohan Douglas, Darrell Duffie, Mark Ferguson, and David Schranz, 2004, Measuring default risk premia from default swap rates and EDFs, *Working paper Cornell University.*

Black, Fisher, and John C. Cox, 1976, Valuing corporate securities: Some effects of bond indenture provisions, *Journal of Finance* 31, 351–67.

Black, Fisher, and M S Scholes, 1973, The pricing of options and corporate liabilities, *Journal of Political Economy* 7, 637–54.

Blanco, Roberto, Simon Brennan, and Ian W. Marsh, 2003, An empirical analysis of the dynamic relationship between investment-grade bonds and credit default swaps, *Bank of England working paper no. 211.*

Campbell, John T., and Glen B. Taksler, 2003, Equity volatility and corporate bond yields, *Journal of Finance* 58, 2321–2349.

Collin-Dufresne, Pierre, and Robert Goldstein, 2001, Do credit spreads reflect stationary leverage ratios?, *Journal of Finance* 56, 1929–1957.

——— , and Spencer Martin, 2001, The determinants of credit spread changes, *Journal of Finance* 56, 2177–2207.

Cremers, Martijn, Joost Driessen, Pascal J. Maenhout, and David Weinbaum, 2004, Individual stock-option prices and credit spreads, *Yale ICF Working Paper No. 04-14.*

Das, Sanjiv, and Rangarajan Sundaram, 1998, A direct approach to arbitrage-free pricing of credit derivatives, *NBER working paper no. 6635*.

Das, S. R., 1995, Credit risk derivatives, *The Journal of Derivatives* 2, 7–23.

de Jong, Frank, and Theo Nijman, 1997, High frequency analysis of lead-lag relationships between financial markets, *Journal of Empirical Finance* 4, 259–277.

Driessen, Joost, 2004, Is default event risk priced in corporate bonds?, *forthcoming Review of Financial Studies*.

Duffee, Gregory, 1999, Estimating the price of default risk, *Review of Financial Studies* 12, 197–226.

Duffee, Gregory R., 1998, The relation between treasury yields and corporate bond yield spreads, *Journal of Finance* 53, 2225–2241.

Duffie, Darrell, and David Lando, 2000, Term structures of credit spreads with incomplete accounting information, *Econometrica* 69, 633–664.

Duffie, Darrell, Lasse Heje Pedersen, and Kenneth J. Singleton, 2003, Modeling sovereign yield spreads: A case study of Russian debt, *Journal of Finance* 58, 119–160.

Duffie, Darrell, and Kenneth Singleton, 2003, *Credit Risk* (Princeton University Press).

Duffie, Darrell, and Kenneth J. Singleton, 1999, Modeling term structures of defaultable bonds, *Review of Financial Studies* 12, 687–720.

Elton, Edwin J., Martin J. Gruber, Deepak Agrawal, and Christopher Mann, 2001, Explaining the rate spread on corporate bonds, *Journal of Finance* 56, 247–277.

Eom, Young Ho, Jean Helwege, and Jing-Zhi Huang, 2004, Structural models of corporate bond pricing: An empirical analysis, *Review of Financial Studies* 17, 499–544.

Fischer, Edwin O., Robert Heinkel, and Josef Zechner, 1989, Dynamic capital structure choice: Theory and tests, *Journal of Finance* 44, 19–40.

Fisher, Lawrence, 1959, Determinants of the risk premiums on corporate bonds, *Journal of Political Economy* 67, 217–37.

François, Pascal, and Erwan Morellec, 2004, Capital structure and asset prices: Some effects of bankruptcy procedures, *Journal of Business* 77, 387–412.

Geske, Robert, 1977, The valuation of corporate securities as compound options, *Journal of Financial and Quantitative Analysis* 12, 541–552.

Houweling, Patrick, Albert Mentink, and Ton Vorst, 2004, Comparing possible proxies of corporate bond liquidity, *forthcoming Journal of Banking and Finance.*

Houweling, Patrick, and Ton Vorst, 2005, Pricing default swaps: Empirical evidence, *forthcoming Journal of International Money and Finance.*

Hull, John, Mirela Predescu, and Alan White, 2004, The relationship between credit default swap spreads, bond yields, and credit rating announcements, *Working paper, University of Toronto.*

Hull, John, and Allan White, 2000, Valuing credit default swaps I: No counterparty default risk, *Journal of Derivatives* 8, 29–40.

Jarrow, Robert A., and Stuart M. Turnbull, 1995, Pricing derivatives on financial securities subject to credit risk, *Journal of Finance* 50, 53–85.

Jones, E. Philip, Scott P. Mason, and Eric Rosenfeld, 1984, Contingent claims analysis of corporate capital structures: An empirical investigation, *Journal of Finance* 39, 611–627.

——— , 1985, *Contingent Claims Valuation of Corporate Liabilities: Theory and Empirical Tests* (University of Chicago Press).

Kim, Joon, Krishna Ramaswamy, and Suresh Sundaresan, 1993, Does default risk in coupons affect the valuation of corporate bonds? a contingent claims model, *Financial Management, Special Issue on Financial Distress,* Autumn, 117–131.

Lando, David, 1997, *Modeling Bonds and Derivatives with Default Risk* . pp. 369–393 (Cambridge University Press).

Ledoit, Olivier, Pedro Santa-Clara, and Michael Wolf, 2003, Flexible multivariate GARCH modeling with an application to international stock markets, *Review of Economics and Statistics* 85, 735–747.

Ledoit, Olivier, and Michael Wolf, 2004, A well-conditioned estimator for large-dimensional covariance matrices, *Journal of Multivariate Analysis* 88, 365–411.

Leland, Hayne, 1994, Risky debt, bond covenants and optimal capital structure, *Journal of Finance* 49, 1213–1252.

——— , and Klaus Bjerre Toft, 1996, Optimal capital structure, endogenous bankruptcy and the term structure of credit spreads, *Journal of Finance* 51, 987–1019.

Leland, Hayne E., 1998, Agency costs, risk management, and capital structure, *Journal of Finance* 53, 1213–1243.

Longstaff, Francis A., Sanjay Mithal, and Eric Neis, 2004, Corporate yield spreads: Default risk or liquidity? new evidence from the credit-default swap market, *forthcoming Journal of Finance*.

Longstaff, Francis A., and Eduardo S. Schwartz, 1995, A simple approach to valuing risky fixed and floating rate debt, *Journal of Finance* 50, 789–819.

Lyden, S., and J. Saranati, 2000, An empirical examination of the classical theory of corporate security valuation, *Barclays Global Investors*.

Martens, Martin, 2003, Estimating unbiased and precise realized covariances, *Working Paper, Econometric Institute, Erasmus University Rotterdam*.

Mella-Barral, Pierre, 1999, The dynamics of default and debt reorganization, *Review of Financial Studies* 12, 535–578.

——— , and William Perraudin, 1997, Strategic debt service, *Journal of Finance* 52, 531–556.

Merton, Robert C, 1974, On the pricing of corporate debt: The risk structure of interest rates, *Journal of Finance* 29, 449–470.

Nielsen, Lars T., Jesús Saá-Requejo, and Pedro Santa-Clara, 1993, Default risk and interest rate risk: The term structure of default spreads, *Working paper, INSEAD*.

Ogden, Joseph P., 1987, Determinants of the ratings and yields on corporate bonds: Tests of contingent claims model, *Journal of Financial Research* 10, 329–339.

Perraudin, William, and Alex Taylor, 2002, Liquidity and the determinants of bond market spreads, *Working paper, Birkbeck College*.

Sharapov, Ilya, 1997, Advances in multigrid optimization methods with applications, Ph.D. thesis UCLA.

Whaley, Robert, 1986, Valuation of american futures options: Theory and empirical tests, *Journal of Finance* 41, 127–150.

Zhou, Chunsheng, 1997, A Jump-Diffusion approach to modeling credit risk and valuing defaultable securities, *Working paper, Federal Reserve Board*.

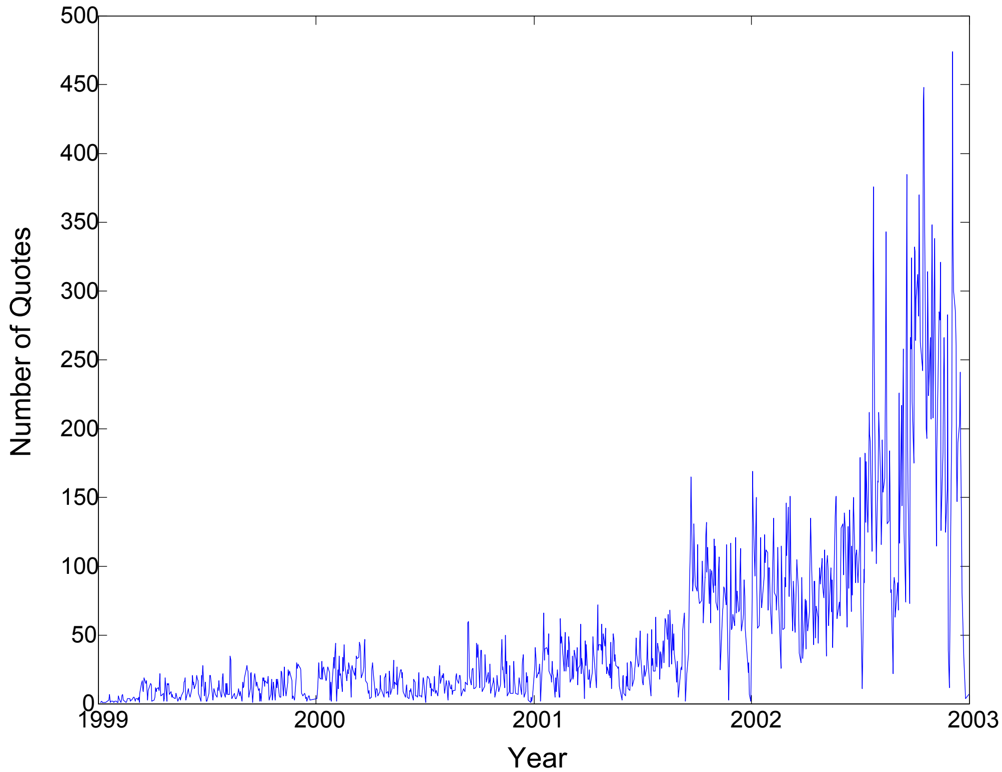


Figure 1: This figure depicts the daily frequency of bid and offer quotes for the CDS premium data during the period January 1999 to December 2002.

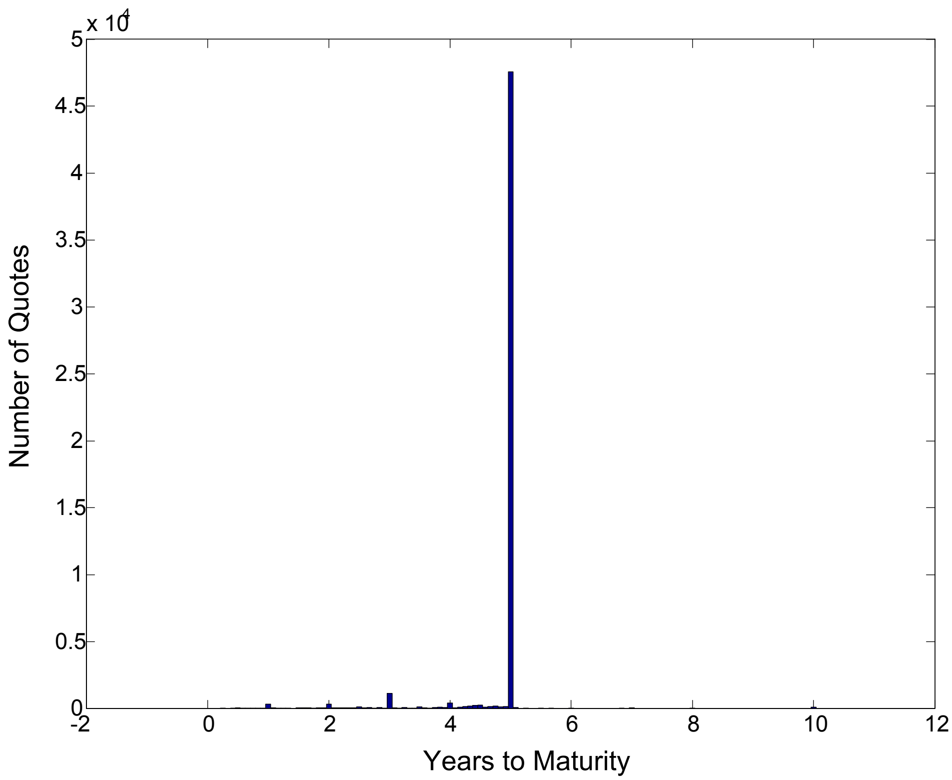


Figure 2: This figure reports a histogram of the maturities of the credit default swaps in our dataset. The figure indicates that the 5 year maturity segment represents the bulk of the market.

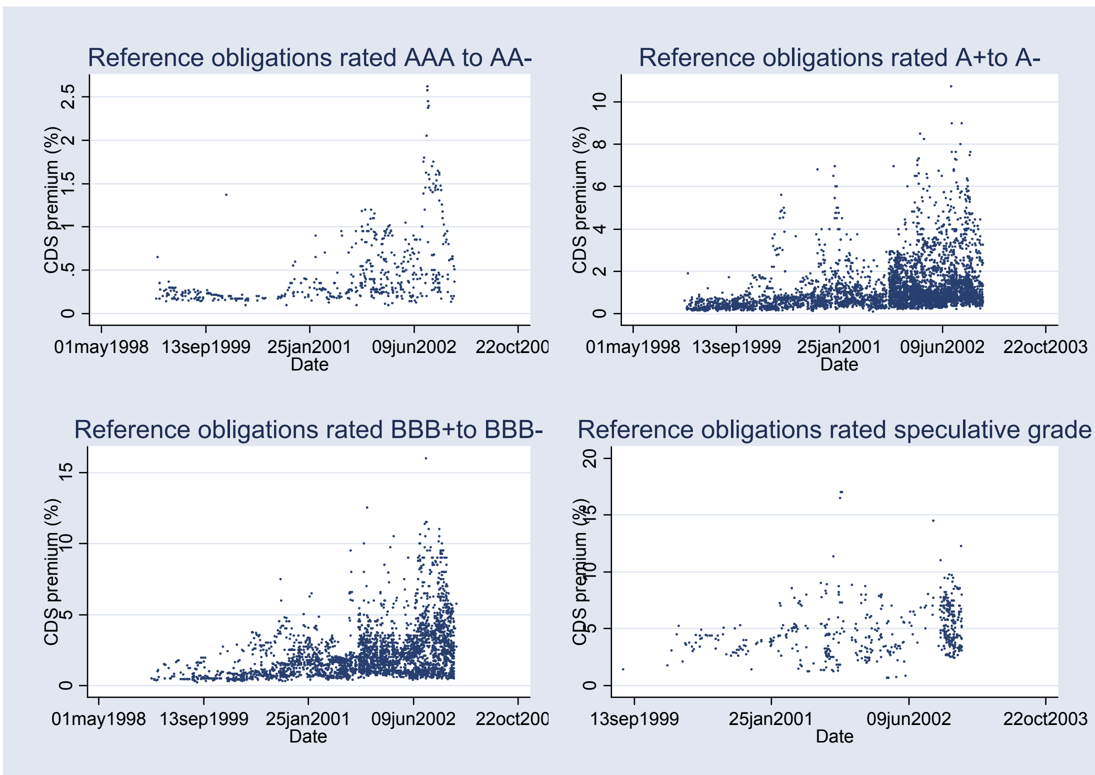


Figure 3: This figure depicts the levels of CDS premia over time and according to rating categories. Data includes bid and offer quotes for all maturities.

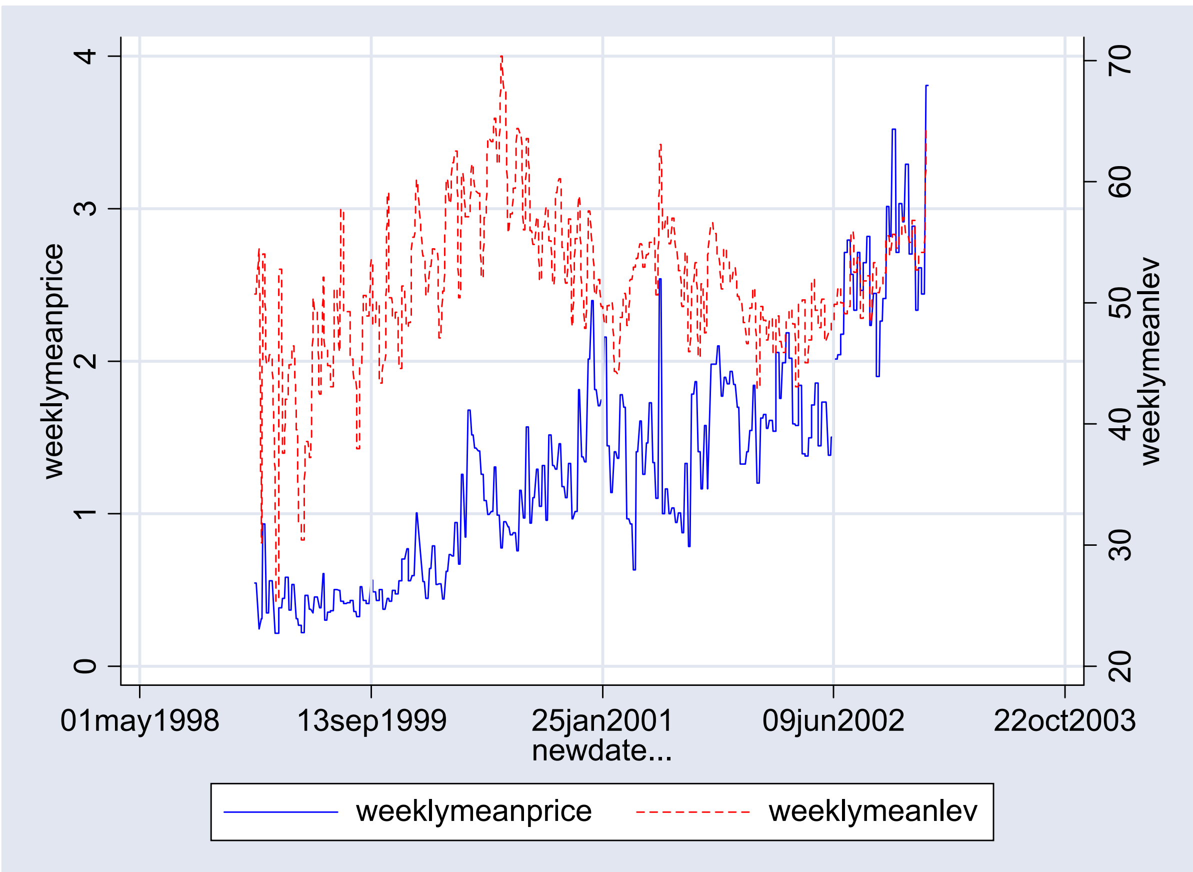


Figure 4: This figure plots CDS premia and firm leverage, both averaged across reference entities on a weekly basis.

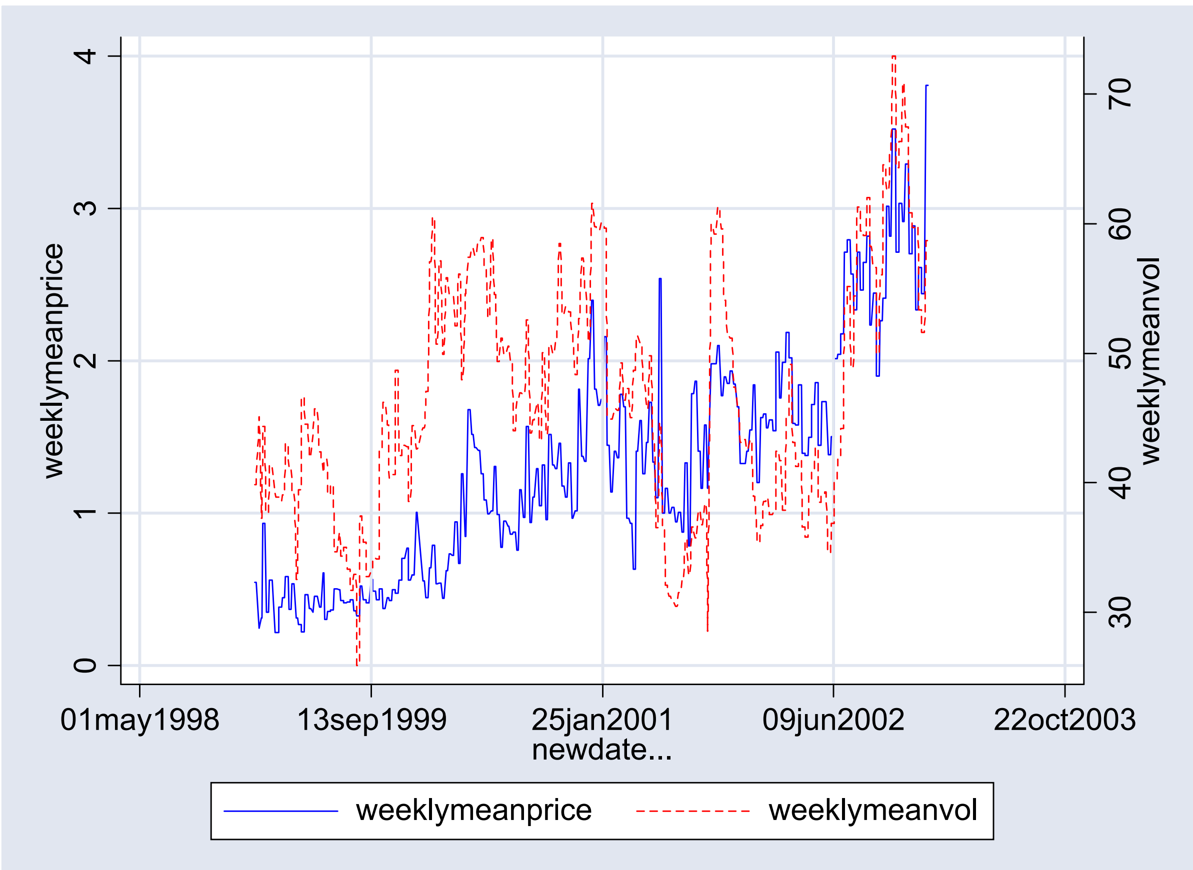


Figure 5: This figure plots CDS premia and equity volatilities, both averaged across reference entities on a weekly basis.

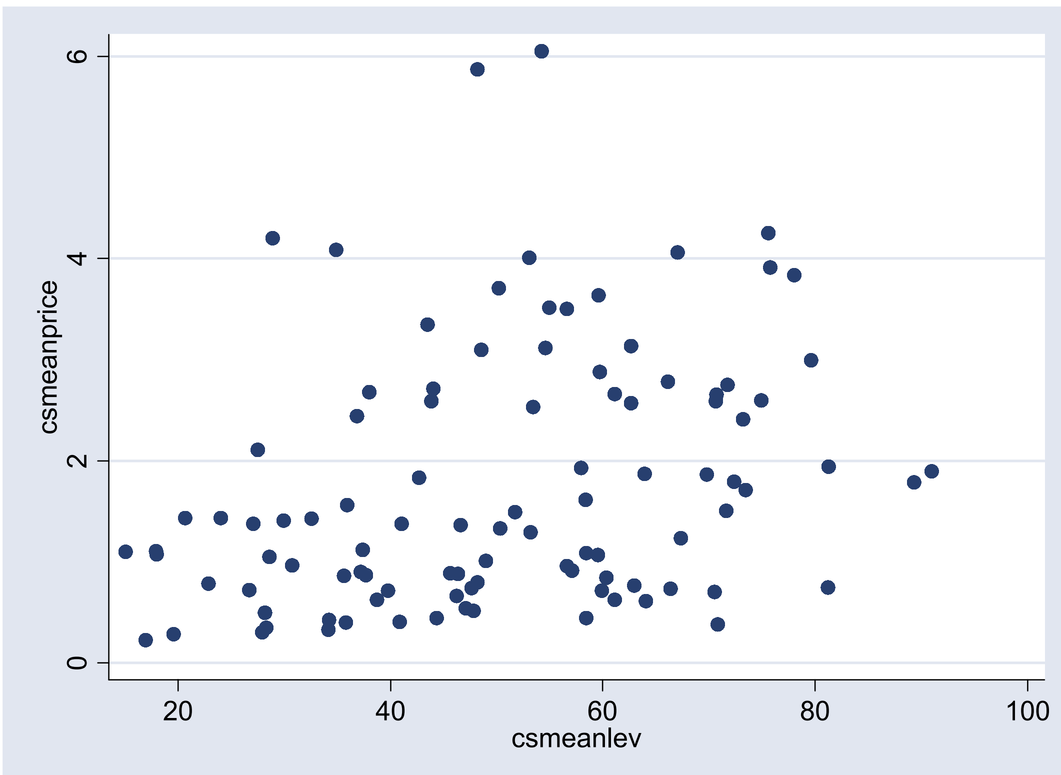


Figure 6: This figure plots the firm-specific (time-series) average of the CDS premia vs. average leverage.

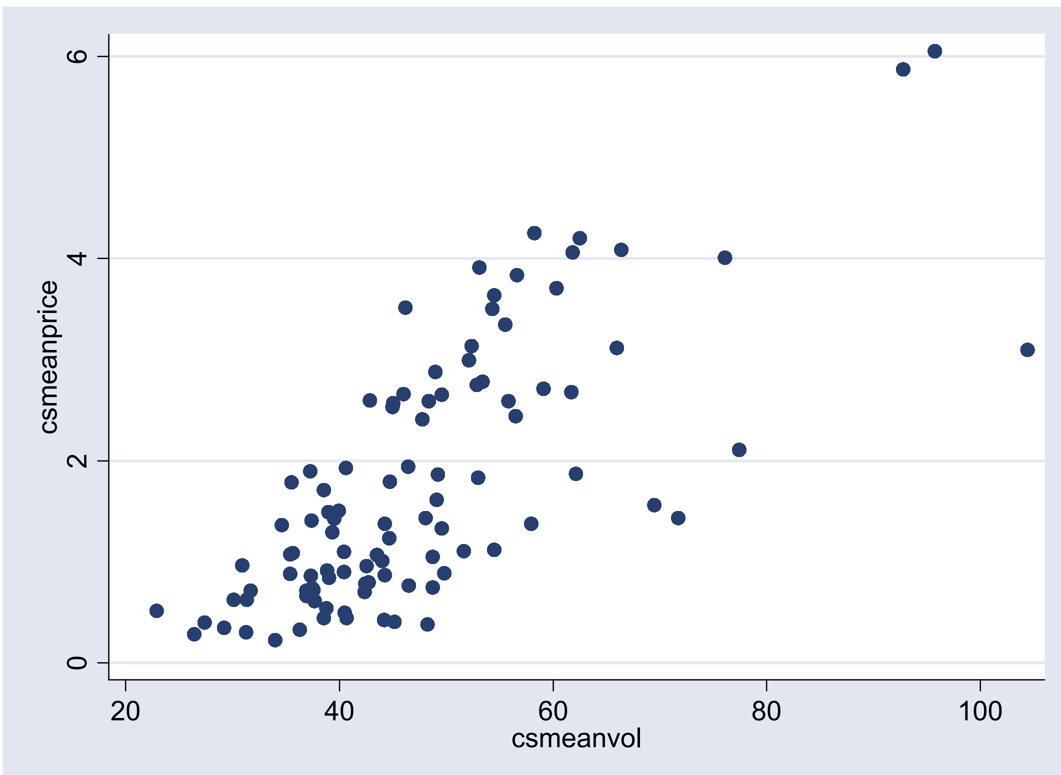


Figure 7: This figure plots the firm-specific (time series) average CDS premia vs. average equity volatility.

Table 1
Summary statistics

This table presents descriptive statistics for the regression variables. It also includes numerical S&P and Moody's credit ratings. Numerical ratings in the sample range from 1 (Aaa) to 20 (Ca) for Moody's and from 1 (AAA) to 25 (in default) for S&P.

	mean	stdev.	5th percentile	95th percentile	Correlation with CDS premium	
					Time series	Cross-sectional
CDS premium (%)	1.80	1.73	0.28	5.30		
Leverage (%)	51.57	17.71	22.75	79.85	0.28	0.23
Volatility (%)	48.80	20.39	25.46	84.09	0.65	0.70
10 year yield (%)	4.92	0.66	3.85	6.11	-0.69	
S&P Rating	7.9	2.1	4	11		
Moody's Rating	8.1	2.2	4	11		
Slope (%)	1.45	0.82	-0.51	2.37	0.59	
2 year yield (%)	3.47	1.39	1.80	6.33	-0.68	
S&P 500	1,111.84	180.87	847.76	1,436.51	-0.70	
Smirk slope (%)	0.59	0.07	0.49	0.70	-0.20	
VIX (%)	29.60	7.19	21.11	43.86	0.52	

Table 2
Regression Using Variables Suggested by Theory

This table presents descriptive statistics and regression results for linear regressions using the three explanatory variables suggested by theory: leverage, volatility and the riskless interest rate. Reported coefficients are averages for regression coefficients from time-series regressions using all observations on a given underlying company. T-statistics are computed based on the time-series regression coefficients as in Collin-Dufresne, Goldstein and Martin (2001).

		Regressions in Differences						Regressions in Levels					
		Bid Quotes			Offer Quotes			Bid Quotes			Offer Quotes		
		Low Rating	High Rating	All	Low Rating	High Rating	All	Low Rating	High Rating	All	Low Rating	High Rating	All
Coefficients	Constant	0.007	0.003	0.005	0.019	0.000	0.010	0.104	-1.072	-0.492	-2.242	-0.783	-1.513
	Leverage	0.072	0.041	0.056	0.060	0.035	0.048	0.076	0.051	0.063	0.100	0.046	0.073
	Equity Volatility	0.011	0.004	0.008	0.023	0.006	0.014	0.017	0.004	0.010	0.023	0.007	0.015
	10-Year Yield	-0.307	-0.118	-0.212	-0.387	-0.169	-0.278	-0.596	-0.100	-0.345	-0.342	-0.057	-0.200
T-stats	Constant	0.87	0.64	1.09	1.56	0.04	1.47	0.09	-1.81	-0.78	-1.66	-1.62	-2.11
	Leverage	6.00	4.82	7.52	4.97	4.85	6.66	5.48	5.86	7.72	6.30	5.69	7.87
	Equity Volatility	4.58	2.97	5.24	5.19	3.61	5.72	3.64	1.97	3.99	3.56	3.39	4.34
	10-Year Yield	-4.49	-2.49	-4.97	-3.13	-2.35	-3.86	-4.27	-1.29	-4.13	-2.28	-0.74	-2.35
R^2		23.3%	21.3%	22.3%	24.2%	23.3%	23.7%	65.5%	57.3%	61.4%	59.6%	52.6%	56.1%
N. of Companies		39	39	78	45	45	90	40	41	81	47	47	94
Avg. N. of Observ.		60.0	59.5	59.7	55.6	61.0	58.3	58.3	60.5	59.4	55.2	60.4	57.8
Avg. Day Btw. Quotes		19.7	19.6	19.7	20.1	19.1	19.6	20.9	19.3	20.1	20.5	19.6	20.1

Table 3
Regression Using Leverage Only

This table presents descriptive statistics and regression results for linear regressions using one of the explanatory variables suggested by theory, leverage. Reported coefficients are averages for regression coefficients from time-series regressions using all observations on a given underlying company. T-statistics are computed based on the time-series regression coefficients as in Collin-Dufresne, Goldstein and Martin (2001).

		Regressions in Differences						Regressions in Levels					
		Bid Quotes			Offer Quotes			Bid Quotes			Offer Quotes		
		Low Rating	High Rating	All	Low Rating	High Rating	All	Low Rating	High Rating	All	Low Rating	High Rating	All
Coefficients	Constant	0.017	0.006	0.012	0.024	0.002	0.013	-5.810	-2.298	-4.032	-7.058	-1.358	-4.208
	Leverage	0.087	0.045	0.066	0.103	0.045	0.074	0.132	0.066	0.099	0.160	0.057	0.109
T-stats	Constant	2.40	1.22	2.65	1.70	0.40	1.73	-4.44	-2.78	-5.11	-4.10	-4.63	-4.59
	Leverage	7.63	5.92	9.14	7.07	7.56	8.82	6.78	5.90	8.44	6.63	6.81	7.88
R^2		14.2%	13.7%	14.0%	12.4%	13.7%	13.0%	44.0%	45.7%	44.8%	40.7%	37.1%	38.9%
N. of Companies		39	39	78	45	45	90	40	41	81	47	47	94
Avg. N. of Observ.		60.0	59.5	59.7	55.6	61.0	58.3	58.3	60.5	59.4	55.2	60.4	57.8
Avg. Day Btw. Quotes		19.7	19.6	19.7	20.1	19.1	19.6	20.9	19.3	20.1	20.5	19.6	20.1

Table 4
Regression Using Equity Volatility Only

This table presents descriptive statistics and regression results for linear regressions using one of the explanatory variables suggested by theory, equity volatility. Reported coefficients are averages for regression coefficients from time-series regressions using all observations on a given underlying company. T-statistics are computed based on the time-series regression coefficients as in Collin-Dufresne, Goldstein and Martin (2001).

		Regressions in Differences						Regressions in Levels					
		Bid Quotes			Offer Quotes			Bid Quotes			Offer Quotes		
		Low Rating	High Rating	All	Low Rating	High Rating	All	Low Rating	High Rating	All	Low Rating	High Rating	All
Coefficients	Constant	0.041	0.018	0.030	0.052	0.015	0.033	0.356	0.120	0.237	0.705	0.156	0.430
	Equity Volatility	0.016	0.007	0.011	0.027	0.010	0.018	0.038	0.016	0.027	0.037	0.017	0.027
T-stats	Constant	5.59	2.47	5.57	3.63	1.92	4.01	0.92	1.02	1.19	1.60	0.95	1.83
	Equity Volatility	6.25	5.25	7.62	6.00	4.94	7.00	5.25	5.01	6.62	4.40	4.73	5.82
R^2		10.1%	6.9%	8.5%	14.4%	11.0%	12.7%	29.7%	24.3%	27.0%	26.9%	23.9%	25.4%
N. of Companies		39	39	78	45	45	90	40	41	81	47	47	94
Avg. N. of Observ.		60.0	59.5	59.7	55.6	61.0	58.3	58.3	60.5	59.4	55.2	60.4	57.8
Avg. Day Btw. Quotes		19.7	19.6	19.7	20.1	19.1	19.6	20.9	19.3	20.1	20.5	19.6	20.1

Table 5
Regression Using 10-Year US Treasury Bond Yields Only

This table presents descriptive statistics and regression results for linear regressions using one of the explanatory variables suggested by theory, the riskless interest rate. Reported coefficients are averages for regression coefficients from time-series regressions using all observations on a given underlying company. T-statistics are computed based on the time-series regression coefficients as in Collin-Dufresne, Goldstein and Martin (2001).

		Regressions in Differences						Regressions in Levels					
		Bid Quotes			Offer Quotes			Bid Quotes			Offer Quotes		
		Low Rating	High Rating	All	Low Rating	High Rating	All	Low Rating	High Rating	All	Low Rating	High Rating	All
Coefficients	Constant	0.030	0.014	0.022	0.036	0.010	0.023	8.848	3.943	6.365	8.608	3.591	6.099
	10-Year Yield	-0.486	-0.285	-0.386	-0.661	-0.356	-0.509	-1.306	-0.596	-0.947	-1.192	-0.500	-0.846
T-stats	Constant	5.20	1.99	4.77	2.60	1.35	2.90	9.12	5.17	9.50	6.81	5.32	8.04
	10-Year Yield	-5.53	-5.79	-7.51	-5.03	-4.63	-6.57	-7.86	-4.57	-8.46	-5.38	-4.24	-6.52
R^2		6.3%	7.5%	6.9%	4.7%	7.9%	6.3%	40.1%	32.6%	36.3%	28.4%	27.9%	28.2%
N. of Companies		39	39	78	45	45	90	40	41	81	47	47	94
Avg. N. of Observ.		60.0	59.5	59.7	55.6	61.0	58.3	58.3	60.5	59.4	55.2	60.4	57.8
Avg. Day Btw. Quotes		19.7	19.6	19.7	20.1	19.1	19.6	20.9	19.3	20.1	20.5	19.6	20.1

Table 6

Regression Using the Regressors from Collin-Dufresne, Goldstein and Martin (2001)

This table presents descriptive statistics and regression results for linear regressions using the benchmark specification in Collin-Dufresne, Goldstein and Martin (2001). Reported coefficients are averages for regression coefficients from time-series regressions using all observations on a given underlying company. T-statistics are computed based on the time-series regression coefficients as in Collin-Dufresne, Goldstein and Martin (2001).

		Regressions in Differences						Regressions in Levels					
		Bid Quotes			Offer Quotes			Bid Quotes			Offer Quotes		
		Low Rating	High Rating	All	Low Rating	High Rating	All	Low Rating	High Rating	All	Low Rating	High Rating	All
Coefficients	Constant	0.014	0.010	0.012	-0.002	-0.007	-0.004	12.046	2.678	7.304	6.201	2.771	4.486
	Leverage	0.063	0.033	0.048	0.059	0.033	0.046	0.075	0.033	0.054	0.073	0.034	0.054
	Equilty Volatility	0.010	0.004	0.007	0.020	0.006	0.013	0.018	0.005	0.012	0.022	0.007	0.015
	2-Year Yield	-0.115	-0.121	-0.118	-0.256	-0.143	-0.200	-1.051	-0.209	-0.625	-0.348	-0.340	-0.344
	Yield Curve Slope	0.005	-0.116	-0.055	0.003	-0.104	-0.050	-0.051	-0.077	-0.064	0.150	-0.020	0.065
	S&P 500	-1.924	-0.284	-1.104	-1.301	-0.034	-0.667	-1.851	-0.411	-1.122	-1.150	-0.453	-0.802
	Smirk Slope	0.144	-0.150	-0.003	-0.524	0.148	-0.188	0.904	0.364	0.631	0.613	0.862	0.738
	Sq. 10-Year Yield	-0.115	-0.117	-0.116	0.009	0.076	0.042	0.114	0.016	0.064	0.013	0.049	0.031
T-stats	Constant	1.12	1.61	1.74	-0.18	-2.23	-0.74	1.64	1.11	1.90	0.96	1.05	1.30
	Leverage	5.28	3.48	6.18	4.22	4.87	5.88	4.79	4.50	6.09	4.64	5.55	6.21
	Equilty Volatility	4.28	2.48	4.81	4.79	3.54	5.48	3.81	2.47	4.36	4.62	3.80	5.51
	2-Year Yield	-0.99	-2.88	-1.93	-1.79	-3.35	-2.67	-1.37	-1.35	-1.62	-0.50	-1.59	-0.95
	Yield Curve Slope	0.03	-1.57	-0.61	0.02	-1.21	-0.49	-0.18	-0.84	-0.43	0.61	-0.20	0.48
	S&P 500	-2.72	-0.90	-2.79	-1.34	-0.15	-1.33	-1.47	-1.11	-1.72	-1.10	-1.08	-1.42
	Smirk Slope	0.26	-0.97	-0.01	-1.04	0.66	-0.68	1.07	1.02	1.39	0.76	2.29	1.66
	Sq. 10-Year Yield	-0.66	-1.48	-1.23	0.04	1.47	0.39	1.59	0.64	1.71	0.19	1.46	0.84
R^2		31.1%	27.9%	29.5%	34.1%	30.5%	32.3%	76.1%	70.6%	73.3%	75.6%	68.6%	72.1%
Number of Companies		39	39	78	45	45	90	40	41	81	47	47	94
Avg. Number of Observ.		60.0	59.4	59.7	55.6	61.0	58.3	58.3	60.4	59.4	55.2	60.4	57.8
Avg. Days Btw. Quotes		19.7	19.6	19.7	20.1	19.1	19.6	20.9	19.3	20.1	20.5	19.6	20.1

Table 7
Panel regressions - offer quotes

This table reports our findings for panel versions of regression (1) and the three univariate regressions (2)-(4). Panel A reports results for OLS regressions with Huber/White/Sandwich variance estimates. Panels B & C report results for regressions with fixed effects and quarter dummies, respectively. The panel contains 5436 offer quotes for 94 different reference entities with at least 25 quotes each.

Panel A

		(1)	(2)	(3)	(4)
Coefficients	Constant	1.118	0.285	-0.608	6.209
	Leverage	0.025	0.031		
	Equity Volatility	0.044		0.052	
	10-Year Yield	-0.532			-0.879
T-stats	Constant	6.310	4.480	-8.030	33.910
	Leverage	21.610	21.910		
	Equity Volatility	26.550		28.890	
	10-Year Yield	-18.430			-25.520
R^2		0.418	0.096	0.326	0.106

Panel B

		(1)	(2)	(3)	(4)
Coefficients	Constant	-1.416	-4.361	-0.015	6.264
	Leverage	0.082	0.123		
	Equity Volatility	0.021		0.039	
	10-Year Yield	-0.376			-0.890
T-stats	Constant	-6.030	-26.810	-0.190	37.890
	Leverage	24.080	37.010		
	Equity Volatility	11.600		22.340	
	10-Year Yield	-13.370			-27.560
R^2		0.698	0.659	0.605	0.572

Panel C

		(1)	(2)	(3)	(4)
Coefficients	Constant	-3.575	-0.841	-1.686	1.876
	Leverage	0.027	0.031		
	Equity Volatility	0.049		0.052	
	10-Year Yield	0.162			-0.261
T-stats	Constant	-7.570	-11.470	-16.880	3.220
	Leverage	23.080	22.520		
	Equity Volatility	26.770		26.190	
	10-Year Yield	1.990			-2.520
R^2		0.460	0.220	0.395	0.130

Table 8
Panel regressions - bid quotes

This table reports our findings for panel versions of regression (1) and the three univariate regressions (2)-(4). Panel A reports results for simple OLS regressions with Huber/White/Sandwich variance estimates. Panels B & C report results for regressions with fixed effects and quarter dummies, respectively. The panel contains 4813 bid quotes for 81 different reference entities with at least 25 quotes each.

Panel A

		(1)	(2)	(3)	(4)
Coefficients	Constant	1.350	0.264	-0.538	6.211
	Leverage	0.023	0.027		
	Equity Volatility	0.039		0.045	
	10-Year Yield	-0.560			-0.917
T-stats	Constant	8.120	4.440	-8.500	34.850
	Leverage	22.330	21.930		
	Equity Volatility	28.280		31.040	
	10-Year Yield	-20.140			-27.080
R^2		0.442	0.088	0.336	0.136

Panel B

		(1)	(2)	(3)	(4)
Coefficients	Constant	-0.733	-4.269	-0.074	6.534
	Leverage	0.074	0.113		
	Equity Volatility	0.016		0.036	
	10-Year Yield	-0.463			-0.983
T-stats	Constant	-2.990	-26.410	-1.150	41.640
	Leverage	21.430	36.020		
	Equity Volatility	11.960		26.120	
	10-Year Yield	-16.660			-31.930
R^2		0.727	0.682	0.611	0.600

Panel C

		(1)	(2)	(3)	(4)
Coefficients	Constant	-3.406	-0.854	-1.347	2.275
	Leverage	0.024	0.026		
	Equity Volatility	0.042		0.043	
	10-Year Yield	0.130			-0.279
T-stats	Constant	-6.830	-10.510	-12.690	3.840
	Leverage	23.750	22.440		
	Equity Volatility	27.910		27.160	
	10-Year Yield	1.660			-2.880
R^2		0.484	0.249	0.420	0.175

Table 9

Principal Component Analysis for Levels using Data on 15 Companies

This table presents results of a principal component analysis using data on the 15 most quoted companies. Principal components is applied either to the levels of the CDS premia or the errors from regression (1) explaining the levels of CDS premia. For each exercise the first two vectors and the percentage of the variance explained by each factor are reported.

Panel A: Bid Levels

Regression errors		Premia	
First Component	Second Component	First Component	Second Component
0.27	-0.15	0.30	-0.01
0.39	-0.01	0.32	0.09
-0.02	-0.49	0.15	-0.38
0.34	0.27	0.31	0.15
0.32	0.03	0.27	0.24
0.20	-0.08	0.25	-0.27
0.04	-0.47	-0.03	-0.51
0.33	-0.01	0.31	0.04
0.00	-0.40	-0.03	-0.36
0.39	0.07	0.27	0.05
0.29	-0.12	0.32	0.15
0.23	-0.31	0.23	-0.33
0.33	0.18	0.32	-0.04
0.05	-0.34	0.17	-0.38
0.08	0.09	0.31	0.15
Explained by PC:			
32.5%	21.6%	58.7%	20.3%

Panel B: Offer Levels

Regression errors		Premia	
First Component	Second Component	First Component	Second Component
0.33	-0.08	0.32	-0.09
0.29	-0.26	0.17	0.36
0.04	-0.44	0.20	0.34
0.21	-0.12	0.32	0.03
0.37	0.05	0.32	-0.08
0.33	0.30	0.31	-0.12
0.35	0.06	0.30	-0.18
0.27	-0.03	0.33	-0.14
0.02	-0.39	-0.08	0.34
0.19	-0.28	0.20	0.37
0.05	-0.47	-0.03	0.48
0.35	0.13	0.28	-0.09
-0.13	0.00	0.27	-0.16
-0.05	-0.37	0.17	0.39
0.38	0.10	0.32	0.07
Explained by PC:			
31.0%	25.1%	55.1%	24.2%

Table 10

Principal Component Analysis for Differences using Data on 15 Companies

This table presents results of a principal component analysis using data on the 15 most quoted companies. Principal components is applied either to the differences of the CDS premia or the errors from regression (5) explaining the differences of CDS premia. For each exercise the first two vectors and the percentage of the variance explained by each factor are reported.

Panel A: Bid Differences

Regression errors		Premia	
First Component	Second Component	First Component	Second Component
-0.47	0.19	0.12	0.53
0.41	-0.14	0.17	-0.50
-0.18	0.35	0.11	0.07
0.15	0.42	0.22	0.07
-0.17	-0.52	0.33	-0.08
-0.32	-0.16	0.18	0.43
-0.06	-0.18	0.35	-0.06
-0.37	-0.09	0.30	0.13
0.17	-0.37	0.32	-0.23
-0.27	-0.04	0.33	0.17
0.04	-0.39	0.27	-0.05
-0.12	-0.07	0.28	0.25
-0.31	-0.11	0.30	-0.23
-0.19	0.07	0.21	-0.21
0.19	0.01	0.22	0.00
Explained by PC:			
24.5%	19.8%	50.2%	18.4%

Panel B: Offer Differences

Regression errors		Premia	
First Component	Second Component	First Component	Second Component
0.41	0.03	0.28	-0.36
0.45	-0.04	0.30	-0.32
0.43	0.04	0.32	-0.24
0.36	-0.33	0.33	-0.08
0.04	0.48	0.26	0.28
0.18	0.10	0.18	0.07
0.37	-0.07	0.19	-0.30
0.23	-0.06	0.30	0.22
0.17	0.26	0.31	0.25
0.06	0.51	0.25	0.43
-0.09	0.28	0.27	-0.04
0.14	0.03	0.21	0.24
0.06	0.13	0.02	-0.04
0.08	-0.04	0.25	-0.31
0.11	0.45	0.21	0.29
Explained by PC:			
30.8%	19.1%	56.5%	15.5%

Table 11
Principal Component Analysis for Levels using Data in Leverage Bins

This table presents results of a principal component analysis using data on all companies grouped in five leverage bins. Principal components is applied either to the levels of the CDS premia or the errors from regression (1) on the levels of CDS premia. For each exercise the first two vectors and the percentage of the variance explained by each factor are reported.

Panel A: Bid Levels

Leverage (%)				**Regressions errors**			**Premia**	
Quintile	From	To		**First Component**	**Second Component**		**First Component**	**Second Component**
1st	17.3	36.8		0.41	-0.08		0.46	0.19
2nd	36.8	47.8		0.48	-0.25		0.48	0.05
3rd	47.8	59.6		0.33	-0.60		0.27	-0.96
4th	59.6	70.1		0.61	0.13		0.49	0.17
5th	70.1	91.0		0.36	0.74		0.49	0.13
				Explained by PC:				
				35.6%	20.8%		68.6%	16.2%

Panel B: Offer Levels

Leverage (%)				**Regression errors**			**Premia**	
Quintile	From	To		**First Component**	**Second Component**		**First Component**	**Second Component**
1st	15.1	34.0		0.24	0.77		0.39	-0.65
2nd	34.0	44.4		0.39	-0.62		0.47	-0.11
3rd	44.4	55.5		0.39	-0.09		0.39	0.70
4th	55.5	65.8		0.60	0.10		0.49	0.21
5th	65.8	81.4		0.52	0.06		0.49	-0.15
				Explained by PC:				
				36.4%	24.0%		66.1%	13.2%