A BETTER AUCTION MECHANISM, AND WHY GOVERNMENTS SHOULD SELL FUTURES RATHER THAN DEBT

Abstract. The mechanism by which governments sell debt at auction imposes unnecessary risks on primary dealers, who charge for these risks by bidding less. A better auction mechanism would increase the price a which governments sell their debt. Further, selling a specially designed futures contract would reduce dealers' balance sheet usage, again increasing revenue, and would also permit the authorities to match cash inflows with cash outflows.

Introduction

Large OECD countries sold just over $1·5 trillion of debt (excluding Treasury-Bills) in 1996, and are forecast to sell only slightly less than this during 1997. British borrowing is only a small share of this, but is still significant: during 1993 the UK authorities borrowed a billion pounds per week, and are now borrowing almost £2 billion per month.

The OECD governments fund most of their borrowing requirements by selling debt at auction. The precise auction mechanism has a substantial effect on the strategies available to potential bidders, and the riskiness of these strategies affects the price that governments receive for their debt. Currently most government issuers accept bids from primary dealers for parts of a large block of debt, and announce within an hour or so the minimum accepted price (or equivalently the maximum yield). Bids above the minimum price are accepted in full, those at the minimum accepted in part, and those below rejected. This widespread and long-used mechanism forces bidders to take risks that a better mechanism would not.

As well as reducing the risks inherent in the auction mechanism, governments can further improve the sale price of their debt by changing what they sell. Currently governments sell debt for settlement within a few days. This implicit linkage between the timing of the auction and the timing of settlement has two disadvantages. First, governments with concentrated debt liabilities must either conduct large infrequent refundings, or alternatively must accept a substantial mismatch between the timing of the receipt of funds and the payment of liabilities. Second, the quick settlement means that primary dealers have little time after the auction to sell debt before it arrives on their balance sheets; and balance sheet usage is expensive for primary dealers.

The unnecessary auction risks, and the awkwardness of timing, slightly depress the sale price of debt, and therefore slightly increase the taxpayers' cost of funds. The price might be only fractionally lower, but saving even a small fraction of £2 billion per month is worthwhile.

In this paper I will first describe how the auction mechanism can be improved, and second I will show that selling futures rather than debt will further help the management of government finances.

Auctions

Let us initially concentrate on the risks inherent in the auction process. There are two.

The first is timing risk: the risk that external events will move markets after bids have been submitted. During the interval between the deadline for bids and the publication of results, a trader's position is unknown and hence unhedgeable. If bond markets start falling after bids have been submitted, then the trader cannot know whether to hedge all, some, or none of the potential exposure. Currently this "timing risk" interval varies from about twenty minutes in France to over two hours in the US.

The second is auction risk: the risk that the auction itself will cause markets to move. If there are few bids at an auction it is likely that any given bid will be filled. The shortage (or low level) of bids becomes apparent when the results are announced, and the market falls in response. So a bid is most likely to be filled if the market is about to fall, and least likely to be filled if the market is about to rally. This "heads you lose, tails is quits" property is a "winner's curse", and bidders charge for this curse by bidding less. Bidders can reduce the winner's curse by hedging their expected holding, but this moves some of the risk onto a "loser's curse", in that an underbidder will find that the hedge causes a loss.

Financial institutions have no objection in principle to risk that is both unquantifiable and unhedgeable, they simply charge for it by bidding less aggressively at auctions. (The central banks and monetary authorities, in their capacity as overseers of the financial system and lenders of last resort, must surely approve of such prudent behaviour.) Neither of these two risks is necessary, and since the public purse ultimately pays for them, they should be removed.

The elimination of timing risk

The timing risk is easy to eliminate: reduce timing risk by conducting auctions quickly. The risk is caused by the delay between the submission of bids and the publication of results,

and can be eliminated simply by eliminating the delay. Dealers should enter their bids into an electronic system, and results of the auction should become available immediately. This has no negative implications for public policy.

The technology is already available: an electronic trading platform (such as LIFFE's APT) could easily be adapted to this role. Indeed, the first exchange to create such a system might find that foreign governments without a financial centre to protect would become keen customers.

The elimination of auction risk

Eliminating the auction risk is much more intricate: reduce auction risk by conducting auctions slowly. The current cause of the auction risk is the absence of price-information flow between participants. By the time price discovery has started it is too late to act on it, as the auction has already finished.

A similar inefficiency can be seen in the daily fixings of LIBOR (London Inter-Bank Offered Rate) by the British Bankers' Association (and indeed, in the interest rate fixings conducted by other bankers' associations in other countries). Each London business day, at 11am, sixteen banks are asked for the rate at which they are able to borrow pounds for 1 month. The highest and lowest four rates are discarded, and the central eight averaged. A similar procedure is used for each of 1-, 2-, …, 11- and 12-month rates in each of thirteen currencies. The averages are then published, as the LIBOR "fixings". These averages often contain anomalies; one or two maturities in a currency having fixings which are cheap or expensive relative to the others. These anomalies can be observed by comparing the fixings with the prices of short-term interest rate derivative contracts, such as futures and FRAs. The anomalies can exist because, by the time the fixings are known, it is too late to trade them. But futures and FRAs are different: trading of these uses live prices, ensuring that information flows between participants and hence that anomalies are traded out.

So the absence of price-information flow during trading impedes market efficiency, and this is as true for auctions as for LIBOR fixings. The solution lies in a mechanism that allows trading and information flow to happen concurrently — as in a secondary market.

This requirement for information flow is inconsistent with the current approach of selling all the paper simultaneously. If instead, an auction of £2 billion (for example) were broken down into forty "auctionettes" of only £50 million each, then earlier auctionettes would provide information for subsequent ones. Furthermore, any individual auctionette would not subject participants to significant risk.

Implementation details

The proposed mechanism requires a number of important details, which have been divided into seven groups. The optimal functioning of the proposed mechanism is quite dependent on these details.

1. Were the auction to be spread over several hours or a whole day, it would be so thin that it might not attract the time and attention of many global investors. To prevent the auction being unduly prolonged, the auctionettes should be one minute apart, with the results for each being published within five seconds.

2. This tough schedule necessitates the elimination of the issuer's discretion from the auction process. The submitted bids should not be reviewed manually, as they are now. Instead, before the auction, the issuer should protect itself against the occasional auctionette receiving only a small number of bids, by publishing in advance a fixed maximum permitted rate of price decline. A maximum rate of price decline equivalent to about one basis point of yield per auctionette would be reasonable. Each auctionette would be subject to a minimum price, namely the previous clearing price minus the maximum rate of decline. The first auctionette would have a just-sub-market minimum price announced one minute before. If an auctionette is not completely sold then those bidding at least the minimum price pay the minimum, and this is deemed to be the clearing price for that auctionette.

   In any auction there might be a small number of auctionettes receiving too few or even no bids. This would be entirely usual, and not a cause for market instability, just a small fall in price and greater trader participation in a subsequent auctionette. These incomplete fills would automatically cause the auction process to be extended by a few minutes.

   It might be that some issuers would also wish to impose a minimum price on the whole auction. This could easily be arranged. The auction would be stopped if bids not below the whole-auction minimum price totalled less than £1 million for ten consecutive

auctionettes. The minimum amount prevents dealers from arbitrarily keeping the auction going by submitting bids for trivially small quantities.

3. Because only one minute separates auctionettes, the software administering the auction would need to allow complicated bidding patterns to be entered or altered quickly. These bidding patterns should include:

   – a bullet bid, such as 99·50 for £50 million;

   – a "ladder" of bids, such as 99·4751, 99·4752, …, 99·5248, 99·5249 and 99·5250 for £100,000 each;

   – repeating the previous set of bids without change;

   – repeating the previous set of bids, but shifted up or down by a specified amount;

   – cancelling all outstanding bids.

   There are exchanges that use sophisticated electronic user interfaces, so these requirements are well within the capability of current trading technology.

4. Most countries currently have an additional rule to prevent squeezes. For example, in the UK no bidder is allowed to take more than 25% of any auction. However, a sequence of auctionettes would ensure that any attempted squeeze would become apparent to the other bidders as the auction proceeded and before the auction ended; they could counter this by bidding more aggressively to acquire paper. Thus this kind of rule would be unnecessary.

   Nonetheless, it might be that a borrower still insists on imposing an upper limit on the quantity of paper acquired by any one dealer. This can be easily accommodated, in one of two ways.

   – The first method is simply to forbid market-makers from bidding for more than a specified share of the auction on their own account, imposing a large financial penalty on transgressors. Immediately after the auction any dealer taking more than the official maximum share of the auction would be obliged to provide the issuer with a list detailing which bids were for their own account and which for clients.

   – The second method would be to create an electronic audit trail. Attached to each bid would be a label, identifying the bidding primary dealer, and specifying whether or not the bid is on behalf of a client. The software would permit vendors to impose constraints based on these labels, possibly more complicated constraints than a simple "share of auction for dealer's own account" rule.

5. Because bid processing would be entirely electronic, there would be no administrative overhead in a very fine price resolution. This would also be desirable, as it would reduce the number of bids subject to scaling down. Bidding resolution currently varies from 0·1¢ (Sweden) through 2¢ (France) and £1/32 (UK) to 5¢ (Netherlands, Belgium and Italy), and from 0·001% of yield (US Treasuries) to 0·01% (Canada, Austria and South African T-bills). A fine bidding resolution has no disadvantages (Swedish traders speak highly of 0·1¢), and so the bidding resolution should be at least as fine as 0·1¢ or an eighth of a thirty-second, these requiring only one extra character of screen space.

6. What information should be revealed after each auctionette? At a minimum:

   – the amount of paper sold in that auctionette;

   – the (uniform) price at which it was sold;

   – the proportion of the lowest accepted bids that were filled (the "scaledown");

   – the amount for sale at the next auctionette;

   – the minimum price of the next auctionette; and

   – the total amount remaining for sale in this auction.

   More information about the number, distribution and size of accepted (or even all) bids could be revealed. But if just the above were supplied, then a trader who had bid for the whole auctionette would know that other traders would receive information that is, except for the price, maximally bullish. This means that a trader who bids for the whole auctionette is reassured that the pattern of other dealers' bids could not cause a fall in the price of the newly acquired debt, so a bidder's risk is further reduced. Indeed, in order to minimise the information released about the distribution of the bids, auctionettes should be uniform-price rather than bid-price; after a uniform-price auctionette it is sufficient to reveal only the lowest accepted price, but after a bid-price auctionette it is necessary to reveal both the lowest accepted and the average accepted price. It is also

important that after the auction the issuer does not reveal any extra information (such as the total quantity of bids received), as this would reintroduce some "auction risk".

There is a slight tweak on this "maximally bullish" reasoning, and it concerns the scaledown. Traders could arbitrarily manipulate the scaledown by bidding for huge quantities of paper at each £50 million auctionette. To prevent this only the top £50 million of bids from each bidder should be used, both for calculating the scaledown and for the subsequent assignment of paper. This would mean that the scaledown factor may now contain information about how many bidders there were sharing the cutoff bid; this information could interfere with the maximally bullish requirement of the previous paragraph. The natural solution is to ensure that the bidding resolution is so fine that even if traders had identical views, it would still be likely that there would be a unique bidder at the lowest accepted price. Thus the bidding resolution should be finer than the typical bid/ask spread in the secondary market divided by the number of primary dealers. In the US, when-issued paper is often quoted with a bid/ask spread of $1/64, and there are almost 40 primary dealers. This suggests that the bidding resolution should be no coarser than 0·05¢ or a sixty-fourth of a thirty-second. As a still finer bidding resolution has no disadvantages, the simplest solution is to allow bidding to four or even five decimal places.

For administrative convenience, as some end-investors are likely to agree to buy paper from primary dealers at the average auction price, this average price should be officially calculated by the auction software and published immediately after the auction. This average would contain no new information, because it could be calculated using only the data published after each auctionette.

7. Some government issuers allow certain classes of investor (such as private individuals, monetary authorities or even primary dealers) to make a "non-competitive bid" for paper, typically subject to an upper limit on the amount of paper. These non-competitive bids are guaranteed to be filled at a price equal to the average of the accepted competitive bids. The new regime would easily accommodate this, provided that non-competitive bids were received in advance of the start of the auction. The total quantity of these non-competitive bids and the quantity thus remaining for sale should be announced just before the auction starts. The quantity of non-competitive bids should be announced before rather than after the auction, so that uncertainty about this information does not become a source of auction risk. After the auction, non-competitive bids would be filled at a price equal to the average sale price in the competitive auction, perhaps rounded to the nearest multiple of the auction's bidding resolution.

During the auction, the quality of price discovery would be very high. Indeed, it may well be that the results of the auctionettes drive the price action in the benchmark bonds and futures, rather than the other way around. Primary dealers would be confident about the price of the new issue to within a small bid/ask spread, and hence, during the auction, could quote fine terms to end-investors.

Furthermore, a similar process could be used by a government in surplus that wished to repurchase debt. Naturally, the "maximum rate of price decline" would become a "maximum rate of price increase", and any minimum price would become a maximum price.

Selling futures rather than debt

As already remarked, there are two difficulties with the timing of auctions.

First, many government issuers either now permit or are about to permit stripping (the breaking of bonds into separately traded cashflows). To maximise liquidity in the coupon strips, bonds should have aligned coupon dates: in the UK, for example, many gilts pay coupons on 7th June and 7th December. This alignment of coupon payments concentrates liabilities into a small number of dates (typically just one or two per year), disrupting their previously even distribution. However, conducting refundings only once or twice a year would concentrate supply so much as to over-burden the debt distribution process, the comparatively large auctions being cleared at an unnecessarily high yield. But the alternative — holding frequent auctions of debt — creates a mismatch between the timing of inflows and the timing of liabilities.

The second difficulty is the increasingly quick settlement following auctions. The Bank of International Settlements has pushed for shorter settlement times, and this means that dealers have less post-auction time in which to offload their debt before it is delivered by the issuer. Many banks are under regulatory pressure not to

enlarge their balance sheet, and this regulatory pressure is reflected in an internal charge for balance-sheet usage imposed on traders. This balance-sheet charge can be another disincentive to bidding.

Both of these problems could be solved, in theory, by delaying settlement after an auction until the next large coupon payment or bond maturity. But in practice, settlement could not be longer than a couple of weeks, as long settlement times increase systemic risk in financial systems, and so government issuers would not wish to be seen encouraging long settlement times. What are the authorities to do?

There is a natural solution, which allows a clean separation of the timing of the transfer of price risk from the timing of the transfer of funds.

Rather than selling debt, a government issuer should sell a futures contract: a contract specially designed for funding. This would work as follows.

The issuer would announce an auction, and the futures exchange would immediately list a new futures contract, called WI ("when-issued"). Such a WI future would be much like a normal bond future, with mark-to-market, initial margin and physical delivery via a clearing house. There would be a single deliverable bond, with a conversion factor of exactly 1, and a single delivery day, so there would be no optionality in the contract. Details of the single deliverable and its delivery day would be specified within the auction announcement, and WI trading would commence immediately after the announcement. As with other financial futures, trading via exchange brokers would be open to all, and those with open positions (including the issuer) would be obliged to provide initial margin, and each day pay or receive variation margin.

WI futures would trade during exchange trading hours until one minute before the auction, at which time all outstanding offers would be cancelled by the software and the auction conducted as described above. During the auction the issuer would be the only seller, and brokers would only accept bids directly from primary dealers. As soon as the auction finishes, usual WI trading would resume and continue until the end of the last trading day. On the delivery day (which would have been specified in the auction announcement), the issuer and any other shorts would deliver the single underlying bond to the clearing house, with money travelling in the opposite direction from those with long positions. The issuer would be provided with a copy of the digital "transcript"

of the auction as soon as possible after the auction.

This use of a futures contract allows the authorities to separate the timing of the transfer of price risk from the timing of the transfer of funds, in a manner that is both transparent and without credit risk. The authorities can supply the price risk associated with debt evenly throughout the year, but can also ensure that receipt of funds matches expected outflows.

Note that it would be easy to enlarge a bond, either by reopening an existing WI contract (if there is one), or by listing a new future with an underlying of an existing bond. Also note that private investors unwilling to engage in mark-to-market could be accommodated: their non-competitive bids would be accompanied by small part-payments, the issuer paying interest on this deemed deposit until delivery day.

Would the WI contracts be liquid?

In recent years, several new styles of futures contract have been attempted, often without success. A futures exchange might well be reluctant to list a new type of contract without some confidence in its success. More concisely: would WI contracts trade?

There is a significant difference between this new contract and most new contracts. Usually, when a new contract is launched the exchange hopes that buyers and sellers would prefer to trade in the pit rather than elsewhere, but there is no certainty that either would do so. But a WI future would have a guaranteed seller who would be committed to selling a known number of contracts. After a £2 billion auction there would definitely be an open interest of at least this nominal value of contracts (possibly less a tiny private-investor allocation). If there were a second tranche of the contract then open interest would be certain to be larger. Whilst the dominant short in the contract (the issuer) would not trade, the primary dealers' long positions will be active and exchanged amongst themselves and with investors.

Indeed, the risk is almost the opposite. The exchange might be concerned that trading in a 10-year WI would undermine an existing 10-year bond future. The neatest way to prevent this would be through careful selection of trading hours. Except on listing, auction and last trading days, WI contracts should commence trading at least thirty minutes after the "usual" contract commences, and there should not be options on WI. However, if a WI contract has an underlying which already existed in cash form, there should be a basis trading facility between the two.

Finally, there is a last improvement to the conduct of auctions that will only reap benefits over the very long term. Currently bids are accepted in confidence, and quite naturally traders do not wish others to know their positions. However, the positions acquired by the primary dealers at auction rarely last long; the primary dealers free balance sheet by on-selling these positions to end-investors. If the full details of the bidding were to be published a year or so after the debt was delivered, then this would be of no immediate consequence to the primary dealers, but would be of great benefit to game-theorists and other academics researching how auctions work in practice. This research might lead to further improvements in the bidding mechanism.

The two improvements described in this paper — splitting auctions into auctionettes and selling futures rather than debt — would cheapen a government's cost of borrowing, slightly reduce the risk in the financial system, and simplify the management of the money-markets. Although either of these improvements would be of substantial benefit by itself, the two complement each other and it is hoped that they will both be implemented.

Julian D. A. Wiseman, 1997