

**Fragmentation and Complementarity:  
The Case of EFPs**

Sharon Brown-Hruska  
School of Management  
George Mason University  
Enterprise Hall MS 5F5  
Fairfax, Virginia 22030  
sbrownhr@gmu.edu

Paul A. Laux  
Weatherhead School of Management  
Case Western Reserve University  
10900 Euclid Avenue  
Cleveland, Ohio 44106  
laux@po.cwru.edu

We are grateful to several Sydney bond and futures market traders for conversations at formative stages of this research. We appreciate valuable comments by Brad Johnson, Gerald Gay, Craig Pirrong, Kasturi Rangan and Marian Turac. We also thank Paul Knapp, Mary Miller and the management of the former Catalyst Institute, which sponsored the research project that inspired this paper.

## **Fragmentation and Complementarity: The Case of EFPs**

### **Abstract**

With the proliferation of alternative markets, concerns have arisen that they may induce lower liquidity on centralized exchanges. In futures markets, the use of an alternative trading mechanism known as exchange of futures for physicals, or “EFPs”, has increased sharply in recent years. EFPs provide a means to obtain futures positions, coupled with offsetting cash positions, away from the centralized exchange. Traders use EFPs to ensure a desired price on complex packages of trades, thus avoiding the transactional risk (slippage) that is inherent in the centralized market. Theoretical analysis establishes that any detrimental effects of fragmenting the centralized market can be offset by traders’ knowledge of another opportunity to trade without transactional risk. If EFPs attract more risk bearing capacity, there could even be a net benefit to the central market. An empirical analysis suggests that EFP trading is motivated by transactional risk, for it represents a larger portion of total trading during periods of unusually high volatility when slippage is apt to be more of a problem. Consistent with the notion that alternative markets can be complementary to centralized exchanges, we find that EFP trading is not associated with reductions in market quality and may act as an outlet for extra volume when markets are under the most stress.

### **1. Introduction**

Securities and derivatives exchanges face substantial challenges from alternative markets. Some have argued that these alternative markets fragment trading, ultimately detracting from price efficiency and liquidity. Proponents of alternative markets generally point to the benefits of increased competition and more cost-effective trading mechanisms. Given that network externalities tend to naturally consolidate trading, alternative markets cannot attract volume without some source of special advantage or opportunity to offer a special benefit to users.

Futures exchanges have argued that alternative markets, such as the over-the-counter derivatives markets, have lower operating costs because they are less regulated. Moreover, since alternative markets often reference prices from central exchanges, the alternative markets have been held to “free-ride” on price discovery. Futures exchanges and regulators have taken strong action to insure the primacy of centralized exchanges. The US Commodity Exchange Act, for example, requires that futures trades occur exclusively on a board of trade, with only very limited exceptions. One such exception is the exchange of futures contracts for physicals, or “EFPs,” which combine cash and futures positions in one transaction agreed to off the exchange floor.

In an EFP, counterparties privately exchange offsetting cash and futures positions at privately agreed-upon terms. Subsequently, the futures leg of the EFP is entered into the futures exchange's normal clearing process; thus some of the terms of the EFP become part of the exchange's trade reporting system. An EFP trade must involve taking simultaneous long and short positions, and so is of the nature of a spread or basis trade. Beyond this requirement, traders have considerable leeway in configuring the various legs of an EFP. In practice, EFPs offer the opportunity for a privately negotiated transaction away from the exchange that nonetheless obtains the benefits of the exchange's clearing process. Depending on exchange rules and technology, other participants might or might not be immediately aware of trading and conditions in the alternative market for negotiated EFPs.

In sharp contrast to on-exchange trading, EFP counterparties can negotiate the terms of an entire package of trades in advance, eliminating "transactional risk" or slippage. Transactional risk arises from the inability of traders to be certain, in advance, of the price at which an order to purchase or sell futures will be consummated. For a complex package, the price differentials between the various cash and futures components are uncertain. Faced with this uncertainty, traders may decide (absent EFP trading) to forgo trades that would otherwise be optimal. All cash and futures legs are transacted in a single package in an EFP, so both counterparties are assured that the expected price relations are realized.

While EFPs are a long-standing minor feature of commodity futures markets, they have recently become very prominent in financial futures markets. EFPs have been found to be especially useful by swaps hedgers, currency basket traders, and bond market traders. This new popularity is explained by the special benefits EFPs provide.

Because of their new prominence as well as their linkage to exchanges' trade reporting processes, EFPs represent a laboratory in which to study possible tradeoffs between the usefulness of alternative trading mechanisms and overall market quality. Alternative markets are often so fragmented or opaque that they are difficult to study empirically. In contrast, because the futures leg of EFPs is cleared through the exchange, EFPs are more observable (at least ex-post).

The major issue investigated in this paper is the tradeoff between the opportunity for useful EFPs and overall market quality on a futures exchange where EFPs are common. To frame the issues and demonstrate the potential tradeoffs, we first present an application of Greenwald and Stein's (1991) transactional risk model to the case of alternative markets. In principle, trading in an alternative market may detriment overall market quality (such as liquidity, volatility and hedging costs). This is so even when the alternative market adds new risk bearing capacity without removing any from the exchange. One potential source of quality reduction is transactional risk in the alternative market, which has spillover effects

on the central market. The problem is that hedgers are apt to strain the risk bearing capacity of the on-exchange market out of concern for transactional risk in the alternative market. However, if the alternative market offers the opportunity to avoid transactional risks, as is the case with EFPs, then the detrimental effects are offset. On-exchange hedgers, knowing the terms of trade in the alternative market, are no longer over-anxious to lay off risk. The beneficial effects of any newly-attracted riskbearing capacity are then dominant.

Whether actual EFP markets are beneficial or harmful depends on various factors identified by the model: the extent to which EFPs are successful in removing transactional risk, the extent of extra riskbearing capacity they attract, and the extent to which the terms of trade in the EFP market are known to on-exchange hedgers. To understand which factors dominate empirically, we study two futures markets in which EFPs have become very popular: the markets for 3 and 10 year Commonwealth Bonds at the Sydney Futures Exchange. Sydney traders have cited the avoidance of transactional risk as a chief motivating factor for using EFPs. Using data from 1992 through 1996, we study the effects of EFP trading on overall market quality, as measured by price volatility, volume, and liquidity. We find that EFP trading does not reduce liquidity. Some evidence indicates that EFP trading is positively related to the level of price volatility, but the causal link flows from volatility and volume to EFP trading, not from EFP trading to volatility. Additionally, the results indicate that EFP trading is statistically generated by past non-EFP trading. This suggests that EFP trading facilities may provide a venue for trades that could not be easily accomodated on-exchange during times of market stress.

We proceed by describing the nature, history, and economics of EFPs in Section 2. Section 3 describes a model of futures market trading in the presence of an alternative market. The model is designed to highlight the importance of some salient characteristics of an EFP market. Section 4 presents our empirical analysis, and Section 5 concludes with a discussion of policy implications.

## **2. Exchange of futures for physicals: Nature, history, and economics**

EFPs were an established commercial practice in agricultural markets as early as the 1920s. They have been used frequently to make or take delivery of a cash commodity while, at the same time, unwinding hedges that had previously been in place.

As an example, consider two traders who could benefit by an EFP. A grain elevator operator might have wheat in storage, hedged by short futures positions. Analogously, a grain merchant might have a commercial need for wheat, hedged by long futures positions. The parties could exchange wheat and liquidate their hedges via an EFP, whereby the merchant

acquires the elevator's wheat and the elevator acquires the merchant's long futures position. In an EFP, this would be accomplished by a single exchange done on privately agreed terms.

Bargaining between the two parties would center on the appropriate futures-cash basis. If the agreed-upon basis at the time of the EFP is negative (e.g., there is backwardation due to a large convenience yield, as is typical in many agricultural markets), then the grain merchant would make a payment to the elevator operator equal to the absolute value of the basis. Wheat delivery would be handled privately, and the futures trade would be entered into the futures exchange clearing system. The previously-existing futures positions are thus offset, and open interest declines.

The advantage of the single EFP trade over separate cash and futures market trades is that the intended futures-cash basis is certain to be realized. The traders eliminate transactional risk. Once the basis for the EFP is pre-established, the actual cash and futures market prices become irrelevant. In contrast, without the EFP, the traders could pre-commit only to the price of the cash trade—the price of the futures trade would be determined in open trading at the futures exchange, possibly obviating the intended basis.

In recent years, EFPs have been extensively used in certain financial futures markets to accomplish a variety of strategies. The reduction of transactional risk is central to all these strategies. In currency futures markets, EFPs are used to trade multi-currency baskets containing both cash and futures contracts as a single package. Such privately-negotiated trades of complex packages are possible because exchange regulations do not typically proscribe the nature of the trade except that it must involve offsetting cash and futures elements in some sense that is often only roughly defined. By the mid-1990s, EFPs accounted for 10 to 15 percent of volume in Chicago Mercantile Exchange (CME) currency contracts.

In swaps markets, EFPs involving precise combinations of cash US Treasury Bonds offset by strips of Eurodollar futures contracts are useful for hedging interest rate swaps. Properly configured, the hedge is a complex spread transaction that offsets the spread position represented by the interest rate swap. The exact configuration of the hedge—number of contracts and maturities—is crucial. Therefore, an EFP is an attractive means of accomplishing the hedge because it removes the transactional risk of price slippage, and it assures that every component of the hedge will be traded.

The use of EFPs for this purpose was sharply curtailed when the CME outlawed the use of EFPs in back-month futures contracts in 1995. The Exchange cited its concern for reduced liquidity and volume in the lightly-traded contracts if volume was drawn away from the exchange. The occasion of its concern was an electronically-facilitated EFP brokerage facility

that was being set up by two institutional brokers, Liberty and Prebon Yemane. Since that decision, the Cantor Futures Exchange has been established; the new exchange has a view to link the futures and cash markets using EFP mechanisms. Also, the CME itself currently offers some products with EFP-like features.

In bond markets, EFPs have a variety of uses. EFPs are a means of accomplishing basis trades in delivery-settled contracts like the Chicago Board of Trade's (CBOT) US Treasury Bond. They are a way of establishing a physical bond position in cash-settled markets like the Sydney Futures Exchange's (SFE) Commonwealth Bond.<sup>1</sup> In contrast to the CME's approach, the CBOT and the SFE have accommodated EFP trading, allowing electronic EFP-brokerage terminals into the exchange.<sup>2</sup> By the mid-1990s, EFPs accounted for 15 to 25 percent of the volume in the SFE government bond contracts. In the huge CBOT treasury bond markets, the proportion of EFP trading was much smaller. Nonetheless, hundreds of thousands of contracts per month were traded via EFPs.

Recently, EFPs have become controversial in the New York metals and energy futures markets. Increased EFP popularity has been blamed for depressed volume and seat prices at the Comex. The Nymex recently obtained regulatory approval to create the "EFS," or exchange for swaps, in the natural gas markets. Its plan has been held up due to members' concerns about the implied increased alternative market trading.

In all these applications, the strong appeal of EFPs is that they can be used to establish complex futures/cash positions with a single transaction. This presents obvious advantages over shopping the various legs of the trade in multiple markets. Avoidance of transactional risk (slippage) is one of the main benefits cited by traders we spoke with as part of our research. In more general discussions, Pagano (1992), O'Hara (1995) and Madhavan (2000) have noted that multiple markets for similar assets can co-exist when the satellite markets address special needs or clientele.

The beneficial or detrimental effects of heavy EFP trading on the futures market overall has not been previously established. Since EFPs fragment trading, there is some *a priori* reason

---

<sup>1</sup> To use an EFP to acquire a cash bond position, a trader would engage in an EFP to obtain a long bond-short futures position, and subsequently liquidate the short futures leg. In conversations, traders claimed that liquidity in the EFP and futures markets was superior to that in the cash bond market. New debt issues are often brought to market this way. Further, since EFPs in Commonwealth bond futures do not need to involve physical Commonwealth Bonds, other sorts of debt issues can also be sold this way. For example, an issue of debt by the New South Wales Treasury was brought out via EFP in Spring 2000 (Hogan, 2000).

<sup>2</sup> The CBOT has also accommodated EFPs via a broad definition of what constitutes an legitimate EFP. A T-bond EFP need not involve any trade in the deliverable bond. The exchange only requires that the physical trade be some asset that is "reasonably correlated" with the T-bond futures.

to expect detrimental effects (as discussed in the introduction). At least some exchange members and managers clearly believe EFPs can be detrimental. On the other hand, Redheffer (1990), in a Commodity Futures Trading Commission (CFTC) memorandum, presents evidence that EFP trading in US Treasury Bonds has been associated with reduced bid-ask spreads. Presumably, this would occur if the availability of EFP basis trading attracts more users to the futures market, resulting in more trading and liquidity overall—similar to the “liquidity externality” that Hendershott and Mendelson (2000) posit for crossing networks that compete with dealer markets for equities.

A special source of concern with EFP trading is transparency, broadly construed to mean the availability to on-exchange traders of information on supply, demand and outcomes in the EFP market. In practice, EFP markets usually exhibit delayed, restricted, or partial trading information to the full market. For example, around the dates of the data used later in this paper, EFPs in SFE contracts were reported with up to a two-and-a-half hour lag. This may be important because some models suggest that the effect of alternative markets on liquidity and price efficiency will depend on information distribution and transparency. For example, when all traders in the market have the same information, Biais (1993) shows that market fragmentation has little effect on trading costs—but when information flow is impeded, transactions costs can rise sharply.

In the sections that follow, we investigate the potential and actual effects on EFP trading on the futures market overall. Based on the facts and reasoning above, the investigation pays special attention to transactional risk and to the flow of information about alternative-market trading to the on-exchange market.

### **3. A model of on-exchange trading with transactional risk and alternative-market trading without transactional risk**

If transactional risk avoidance is at the center of EFP trading, then its potential effects for on-exchange futures markets can be understood by working out the effects of a no-transactional risk alternative market that coexists with an on-exchange market with transactional risk. This section provides such an analysis by applying Greenwald and Stein’s (1991) microstructure model to the present situation.

Greenwald and Stein (1991) focused on a single market, comparing market-quality outcomes with and without transactional risk. With some adjustments, that model can be used to study the case of a fragmented market. Here, the “on-exchange” market is subject to the transactional risk typical of open-outcry trading, while traders in the “alternative” market may be able to pre-contract to avoid transactional risk. When such pre-contracting is possible, we identify the alternative market with EFP trading.

In the most basic EFP situation, one group of traders holds a combined long cash-short futures situation. That first group of traders perceives either business or risk-shifting incentives to sell the combined position to other traders.<sup>3</sup> We refer to the first group of traders as sellers or hedgers. The second group of traders may be buyers who require the cash market positions for business purposes, or else are speculators. In our model, we imagine them to be speculators who will transact only if they can earn a risk premium. With this convention, we can measure liquidity and hedging costs from the point of view of the sellers.

### 3.1 Model timeline and market participants

The model includes up to two trading opportunities for hedgers/sellers to lay off positions to speculators. The model accommodates cases with and without fragmentation, and with and without transactional risk. In our model, transactional risk is always present at a “first” trading opportunity, which we identify with on-exchange trading of the futures component of the EFP and exchange of the cash component. In the case where the market is unfragmented, there are no further trading opportunities.

When present, trading at a “second” opportunity, the alternative market, is also subject to transactional risk in a base case. However, in our focal case, which we identify with EFPs, the second market is not subject to transactional risk. We compare volume, liquidity, and hedging costs in this fragmented, reduced-transactional risk case with the outcomes for fragmented and unfragmented markets in which trades at all dates are subject to transactional risk.

The model employs a four-date architecture, with times  $t = \{0, 1, 2, 3\}$ . Endowments and market rules are established at the first time,  $t = 0$ ; Trading takes place at times 1 and (possibly) 2. Time 1 is the on-exchange trading opportunity. Time 2, which is considered to be only slightly after time 1, is the alternative-market trading opportunity.<sup>4</sup> Resolution of all uncertainty and settling-up occurs at the final time,  $t = 3$ . Anyone holding a long position at time 3 receives a payoff of  $v$ , an innovation whose value is determined just before time 3.<sup>5</sup> The innovation is a mean-zero normal random variable with variance  $\sigma_v^2$ .

---

<sup>3</sup> Though the traders hold a hedged position, they face basis risk because the futures and the cash positions in an EFP are not generally perfectly correlated.

<sup>4</sup> To understand the characteristics of an unfragmented market, we present some benchmark results below for the case where time 2 trading is eliminated altogether.

<sup>5</sup> The assumption that there is no value innovation between time 0 and time 1 is merely a simplifying assumption and is not important for the results. The assumption that there is no value innovation between time 1 and time 2 reflects the idea that the associated trading opportunities are very close in time, so that little or no news arrives in between.

There are three groups of traders. First are  $k$  hedgers/sellers endowed with an aggregate long position of  $s$  risky units. The endowment is one source of price uncertainty in the model, for it is not known to other traders. From their point of view,  $s$  has mean  $\mu_s$  and unconditional variance  $\sigma_s^2$  as of  $t = 0$ .

Hedgers have risk-averse utility functions of terminal wealth with a coefficient of absolute risk aversion equal to one.<sup>6</sup> Their motive for trade is to lay off risk to other traders. They will optimally pay a price-concession to achieve this risk-sharing benefit. Because the hedgers trade in response to their risk endowment, the size of the endowment will be effectively revealed by prices during trading. Price discovery in this model means that expressed trading desires result in a market price that reveal the value of  $s$ .

In addition to the hedgers, there is a group of  $n_1$  speculators/buyers in the on-exchange market at time 1, and also a separate group of  $n_2$  speculators/buyers in the alternative market at time 2. Like the hedgers, speculators have risk-averse utility functions of terminal wealth with a coefficient of absolute risk aversion equal to one. However, unlike the hedgers, speculators begin with no endowment. They are therefore willing to take on a position in the hope of a profit, provided the expected price is right. To allow for the possibility of transactional risk in the model, speculators are assumed to place market orders to buy; that is, they place their orders before observing the fill price.

Of the EFP situations described in Section 2, the use of EFPs to bring new bond issues to market at the Sydney Futures Exchange is closest to our model. The sellers (hedgers) in our model correspond to the bond-issuers or underwriters, in that they begin with a risky endowment which they desire to transfer. The buyers (speculators) in our model correspond to the investors who purchase the bond issue via an EFP.<sup>7</sup>

### 3.2 Transactional Risk

A precise definition of transactional risk is required for modeling purposes. In this context, transactional risk means that some traders must specify their orders without knowing the equilibrium price at which trading will occur. Thus, they face the likelihood of price slippage between the price they expect and the price they realize.

Consider speculators, the risk bearers-for-pay in our model, who place market orders to buy. Because they face transactional risk in the on-exchange market at  $t = 1$ , they shy away from

---

<sup>6</sup> The assumption that risk aversion equals unity is used to simplify the expressions that follow; it does not affect our conclusions.

<sup>7</sup> Our model does not treat the eventual offset of the futures leg of the EFP by the investors, but only the situation at the time of the offering.

trading at that date. The economic source of this transactional risk is that each speculator must express his demand before becoming fully informed about the presence of other speculators willing to take on risk. Therefore, speculators can only estimate the market clearing price. There is likely to be some slippage between the price that speculators expect and the price they obtain. Formally,  $n_1$  and  $n_2$  are ex ante normal random variables for the spectators, where  $n_t$  has mean  $\mu_t$  and unconditional variance  $\sigma_t^2$ .

In contrast, in the case we identify with EFPs, speculators do not face transactional risk in the alternative market at time 2. In that case, this is accomplished by allowing the time 2 speculators to condition their demand for contracts on the actual population,  $n_2$ . We include a group of traders who do not directly face transactional risk so that we can show the indirect but substantial market-wide effects when other traders continue to face such risks.

Hedgers at each trading opportunity know how many of them are present. They place their sell orders after the speculators express their demands. Thus, unlike the speculators, hedgers do not face direct transactional risk at all—they can condition their sell orders on the equilibrium price. Nonetheless, as we will show, the detrimental effects of the speculators' transactional risk are transmitted to them.

### 3.3 Equilibrium characteristics of an unfragmented on-exchange market

We first consider a market in which there is no trading at time 2, i.e., no alternative market at all. Only the unfragmented on-exchange market at time 1 exists in this benchmark case, which we use to establish a basis for comparison to fragmented situations with an added alternative market. Notationally, solution quantities for this particular case are designated with a subscript  $u$ , for “unfragmented.”

The solution method for this model is standard, consisting of finding the competitive price that is consistent with the optimal supply and demand functions of the traders. Greenwald and Stein (1991) outline the solution method. A representative hedger's optimal holdings,  $h_{1u}$ , after an equilibrium volume of trade at date 1, are given by

$$h_{1u} = \frac{-P_{1u}}{\sigma_v^2}, \quad (1)$$

where

$$P_{1u} = -\frac{(s - n_1 d_1) \sigma_v^2}{k} \quad (2)$$

is the price at which time 1 trades occur.<sup>8</sup> In (2),  $d_1$  refers to the optimal demand of a representative speculator, calculated on the conjecture that all speculators behave the same way. In the unfragmented market with transactional risk, this demand is given by

$$d_{1u} = \frac{\mu_s}{k + \mu_1 + \frac{\sigma_v^2(\sigma_s^2 + \sigma_1^2 d_{1u}^2)}{k}}. \quad (3)$$

The speculator's demand in (3) depends only on parameters known before the price is established. This is the essence of the transactional risk: demands cannot be conditioned on the actual price,  $P_{1u}$ , but only on the expected price (given by (2), but with  $s$  and  $n_1$  replaced by their expected values). As in Greenwald and Stein (1991), trading exerts an externality. The representative speculator's demand is depressed when other speculators trade heavily, as evidenced by the presence of  $d_{1u}^2$  in the denominator of (3). Heavy trading in general increases the variance of the price at which any speculator will trade, and so they all back away from the market.<sup>9</sup> Importantly, depressed speculator demand due to the transactional-risk externality also implies depressed market volume—for market volume is simply the number of speculators times the representative speculator's demand.

Even though hedgers do not face transactional risk directly (i.e., they can condition their orders on the return price), there are several indications that they still bear its ill effects. First, the marginal willingness of the representative speculator to accommodate an additional risky unit in the marketplace, defined as the derivative of  $d_{1u}$  with respect to  $\mu_s$  as calculated from (3), is smaller when there is transactional risk (because of the presence of  $d_{1u}$  in the denominator). Second, and related to the first point, the ex-ante inverse depth measure that hedgers expect is more negative because of the transactional risk. Inverse depth, which indicates the extent to which price must decline to induce speculators to accommodate another unit of expected supply, is defined as the time-zero expectation of the derivative of  $P_1$  with respect to  $\mu_s$ :

$$E_0 \left( \frac{dP_{1u}}{d\mu_s} \right) = \frac{\sigma_v^2}{k} \left( \frac{\mu_1}{\mu_1 + k + \sigma_v^2(\sigma_s^2 + \sigma_1^2 d_{1u}^2)} - 1 \right) < 0. \quad (4)$$

In this context, where hedgers are sellers, a negative inverse depth measure implies that hedgers expect to sell at less advantageous prices as  $\mu_s$ , their total risky endowment, is

---

<sup>8</sup> The price is negative because it consists of only a risk premium, which the hedgers must pay to the speculators to induce the trade.

<sup>9</sup> Mechanically, the final term in parenthesis in the denominator of (3) is proportional to the price variance.

larger. In the absence of transactional risk, the term  $\sigma_v^2(\sigma_s^2 + \sigma_1^2 d_1^2)$  would vanish, implying better terms of trade for the marginal hedger.

Finally, an indicator of the hedger's expected hedging cost,  $EHC_u$ , defined as the expectation at date 0 of the squared deviation of the market price from the terminal expected value, is larger because of the externality.  $EHC_u$  is given by:

$$EHC_u = \sigma_v^2 + \left( \sigma_v^2 \frac{s - \mu_1 d_{1u}}{k} \right)^2 + \frac{\sigma_1^2 \sigma_v^4 d_{1u}^2}{k^2}. \quad (5)$$

In the absence of transactional risk, the final term would be eliminated, reflecting a reduction in expected hedging cost.

### 3.4 Equilibria in a fragmented market with transactional risk

If the benchmark on-exchange market above becomes fragmented by the addition of another trading opportunity that is also subject to transactional risk, things only become worse. We present this case for contrast to the subsequent EFP-like case where transactional risk is eliminated at the second trading opportunity.

#### 3.4.1 The second trading opportunity

With two trading opportunities, the model is solved by backward induction, starting with the time 2 trading opportunity. Traders optimally choose their demand (in the case of speculators) or supply (in the case of hedgers) to maximize the expected utility of terminal wealth. Because all traders have constant absolute risk aversion, the solutions follow a standard pattern. The procedure is outlined in Greenwald and Stein (1991). Notationally, we distinguish outcomes for this case with a subscript  $f$ , for “fragmented.”

A representative hedger is optimally willing to hold  $h_{2f}$  contracts from time 2 to time 3, where  $h_{2f}$  is given by

$$h_{2f} = \frac{-P_{2f}}{\sigma_v^2}, \quad (6)$$

and  $P_{2f}$  is the price at which time 2 trades occur. Considering that hedgers must hold their endowment less whatever has been purchased by speculators, the equilibrium price at time 2 is

$$P_{2f} = -\frac{s - n_1 d_{1f} - n_2 d_{2f}}{k}. \quad (7)$$

where  $d_{1t}$  represents the demand for contracts by a representative speculator at date  $t$ . These expressions for the time 2 market are very similar to the corresponding expressions for speculator holdings and price, (1) and (2), in the one-shot on-exchange market.

The representative speculator's demand,  $d_{2t}$ , is

$$d_{2f} = \frac{s - n_1 d_{1f}}{k + \mu_2 + \frac{\sigma_1^2 \sigma_v^2 d_{2f}^2}{k}}. \quad (8)$$

Similar to the one-shot market, representative speculator demand is depressed by an externality: heavy demand by other speculators drives up the variance of the price they will face, so the risk-averse representative speculator backs away from the market.

Other than being shifted one date later, the main difference between (8) and speculator demand in the one-shot on-exchange market is that speculators condition their demands on the actual aggregate hedger endowment, for, as is shown below, this has previously been revealed in on-exchange trading. Thus, (8) does not involve the mean and variance of that endowment. In this sense, the alternative market free-rides on the price discovery provided by the on-exchange market.

### 3.4.2 The harmful indirect effects of time 2 transactional risk for the time 1 on-exchange market

The transactional-risk externality in the time 2 alternative market indirectly harms the earlier on-exchange market. The price hedgers' receive is lower, depth is impaired, and expected hedging cost is greater. Additionally, on-exchange volume is most likely lighter simply because of the fragmentation of available risk bearing capacity.

To focus most cleanly on the indirect deleterious effects of transactional risk, we consider the case where all time 2 speculators represent additional capacity, so that the distribution of  $n_1$  is unaffected by the fragmentation. In particular, this means that no speculative capacity is drawn away from the on-exchange market to the alternative market.

In this case, the equilibrium price for the on-exchange market is

$$P_{1f} = P_{1u} + \sigma_v^2 \frac{\mu_2 d_{2f}}{k} - \sigma_v^2 \frac{s - n_1 d_{1f}}{k} \frac{\sigma_1^2 \sigma_v^2 d_{2f}^2}{k^2}, \quad (9)$$

where  $d_{1t}$  is the representative speculator's demand at time 1.

The expression shows that  $P_{1f}$  differs from  $P_{1u}$  (from (2)) for two reasons. First, the term involving  $\mu_2$  is due to the presence of additional speculators in the time 2 alternative market; hedgers do not have to pay so large a risk premium (i.e., accept such a low price) on-exchange because there are additional risk bearers available. If the fragmented market brings little additional risk bearing capacity, then this benefit would be small.

The final term shows the harmful effects of alternative-market transactional risk. The on-exchange price is less when alternative-market trading is heavy, for alternative market trading increases the variance of the price speculators will face in that market.<sup>10</sup> As a result, hedgers know that alternative-market speculators will be more hesitant to trade. This, in turn, makes hedgers more eager to lay off risk in the on-exchange market, depressing the equilibrium price.

Note that equilibrium speculator demand in the fragmented on-exchange market will be greater than (less than) in the unfragmented market depending on whether  $P_{1f}$  is less than (greater than)  $P_{1u}$ . This is because time 1 speculators trade only once, and their demand is predicated only upon the premium they receive for risk-bearing. By reference to (9), then, it is clear that volume in the on-exchange market may be greater or less than in the unfragmented case. It would be greater in the unfragmented case if the volume-depressing effect of increased risk-bearers in the alternative market is swamped by the volume-increasing influence of hedgers being more anxious to trade.

Applying these conclusions to the situation of futures exchanges, it is easy to see why exchanges would complain about alternative markets in general. On-exchange volume is apt to fall if the alternative market attracts significant new risk-bearing capacity, making it more difficult for the exchange to cover fixed costs. The volume effect could be offset by the spillover effect of alternative-market transactional risk. But, in that case, on-exchange hedgers perceive artificially elevated levels of price risk and unnecessarily strain the risk-bearing capacity of speculators.

With a transactional-risky alternative market, on-exchange market depth is impaired, irrespective of whether volume increases, as evidenced by a more-negative inverse depth measure:

$$E_0 \left( \frac{dP_{1f^*}}{d\mu_s} \right) = E_0 \left( \frac{dP_{1u}}{d\mu_s} \right) - \frac{\sigma_v^2}{k} \frac{\sigma_1^2 \sigma_v^2 d_{2f^*}^2}{k^2}. \quad (10)$$

---

<sup>10</sup> The ratio at the end of (9) is, in fact, the expression for the variance of the price in the alternative market.

This measure of the price concession needed to induce speculators to accommodate another unit of endowment risk in the on-exchange market is greater than in the unfragmented case, commensurate with the strain due to the indirect effects of transactional risk. Similarly, transactional risk in the alternative market increases overall expected hedging costs.  $EHC_e$ , the analogue to  $EHC_u$  from (4), can be shown to be larger by terms that depend directly on the intensity of trading in the alternative market.

In the comparisons above, on-exchange risk bearing capacity was held constant. If some on-exchange risk bearing capacity defects to the time 2 alternative-market, the on-exchange results would be even less favorable than those reported.

### 3.4.3 The beneficial effects of eliminating transactional risk in the time 2 alternative market: A case like EFPs

The fragmented market may have more total risk bearing capacity than the one-shot on-exchange market if the additional trading venue can attract additional risk bearing capacity. Additional capacity is beneficial to hedgers, all else equal. Is there a way to allow for the additional trading venue without inducing the ill effects of alternative market transactional risk for the on-exchange market? Yes—if the alternative market trading mechanism can be arranged to eliminate transactional risk. Since this is exactly what EFPs are reputed to accomplish, the case is directly relevant.

In this section, we consider a variation of the model where  $n_2$ , the number of risk bearers in the alternative market, is known in advance to all agents. Notationally, we designate the outcomes for this case with an “e,” since it represents a market with EFP-style no-slippage trading. With this information, the time 2 price can be calculated as soon as the time 1 on-exchange market has met. Thus, there is no transactional risk in the alternative market.

Because conditions in the alternative market can be anticipated, the markets are not economically separate. Though the alternative market still rides on the price discovery benefit of the on-exchange market, it now pays for the ride by providing the on-exchange market with an outlet for excess volume on terms that are known to all. On-exchange speculators do not then strain the risk bearing capacity of the on-exchange market out of concern for the unexpectedly poor price they might face in the alternative market, and the terms of trade are more beneficial to hedgers.

Specifically, in the case where the number of alternative-market speculators,  $n_2$ , is known in advance to all, the representative speculator’s time 2 demand is

$$d_{2e} = \frac{s - n_1 d_{1e}}{k + n_2}, \quad (11)$$

where  $n_1 d_{1e}$  is volume in the time 1 on-exchange market. By comparison to (8) on the assumption that  $n_2$  is at its mean ( $\mu_2$ ), it is evident that alternative-market volume tends to be larger when transactional risk is eliminated. The absence of any terms involving the trading of other speculators in the denominator of this expression indicates that the transactional-risk externality has been removed.

Hedgers are better off than in the fragmented case, and as well off as in the unfragmented case. The equilibrium price, which is the same for the time 1 and time 2 markets because both equilibria are established with exactly the same information available, is

$$P_{1e} = P_{2e} = -\sigma_v^2 \frac{s - n_1 d_{1e} - n_2 d_{2e}}{k} . \quad (12)$$

On the assumption of equal trading volume in each case, it can be shown that the expression for on-exchange market depth is exactly the same as for the unfragmented market (given in (4)) and is superior to that for the fragmented market (given in (10)). Similarly, expected hedging costs in the on-exchange market are no worse than in the unfragmented market, and are superior (i.e., less) to the extent that the alternative market adds any new risk bearing capacity to the marketplace.

On-exchange speculators still face transactional risk in their own market. Volume is depressed as a result (very similar to the unfragmented market case in (3)):

$$d_{1e} = \frac{\mu_S - n_2 d_{2e}}{k + \mu_1 + \frac{\sigma_v^2 (\sigma_S^2 + \sigma_1^2 d_{1e}^2)}{k}} . \quad (13)$$

The comparison of (13) to (3) shows that, even if no on-exchange speculators defect to the alternative market, the two markets now share in the hedgers' order flow. If volume *per se* generates fees, etc. to help the exchange cover its fixed costs (not part of the model), this may be a source of some exchanges' complaints. Relevant fixed costs may include the costs of maintaining a trading floor that engenders the price discovery, as well as the costs of maintaining margining, clearing, and settlement systems. Without these services, the alternative market might not be viable. One possible solution is to charge fees for EFP trading that are sufficiently large to compensate.

#### 4. Empirical examination of the economic effects of EFP trading in the Sydney Futures Exchange bond markets

The section analyzes EFP trading in Sydney Futures Exchange (SFE) Commonwealth Bond futures contracts. The preceding theoretical analysis establishes that, while alternative-market trading can have a detrimental effect for the on-exchange market, this need not be

the case. Indeed, alternative-market EFP trading that eliminates transactional risk and thereby attracts new risk bearing capacity could benefit the on-exchange market. In view of those findings, the purpose of this analysis is to establish the effects of EFP trading using data from a market where EFP trading has been a significant factor. SFE 3 and 10 year Commonwealth bond futures experience a high volume of EFP trading, facilitated by screen brokerage systems.

The prevalence of EFP trading in Commonwealth bond futures markets suggests that these markets are useful laboratories to examine the potential effects of EFPs on underlying futures markets. Additional factors also make the SFE contracts a good testing ground. Cash bond traders in Sydney have suggested to us that EFPs offer advantages in terms of enhanced liquidity, lower transactions costs, greater flexibility in timing, and even greater anonymity. The fact that even relatively small trades are done via EFPs in the Commonwealth Bond market suggests that these benefits may be even more prominent than in other markets that use EFPs. Traders we spoke with maintain that the availability of EFPs has expanded the futures markets; cash bond traders who had not previously used futures are participating in EFPs and subsequently unwinding their open positions via the futures pits. This is consistent with the possibility that EFP trading may have added risk bearing capacity, as suggested in the previous section.

During the sample period, exchange regulations required reporting of trades within two and a half hours of execution, but this information was not electronically disseminated to the exchange floor. The model in the previous section assumes that on-exchange traders can observe the terms of trade in the alternative market; in fact, this information is crucial, since it conditions traders' on-exchange demand. But since this was not actually the case in Sydney, the possibility arises that the Exchange might bear the costs of fragmentation without the positive spill-over benefit of reduced transactional risk in the alternative market.

#### **4.1 Data and descriptive statistics**

The data used in the analysis, provided by the SFE, consist of the intraday trading record on the 3 year and 10 year Australian Commonwealth Bond futures over a period from January 4, 1993, through April 30, 1996, as well as daily data on total futures contract volume beginning with 1988. Additionally, we have daily data on EFP trading for January 2, 1992 through April 30, 1996.

During the period of the data, the SFE Commonwealth Bond contracts each call for physical delivery of a 12 percent coupon bond with A\$100,000 face value. As in other bond futures markets, conversion factors are available to accommodate the cheapest-to-deliver bond. Contracts expire quarterly in March, June, September and December. According to

conversations with traders, EFPs are typically used in these markets to facilitate the taking and closing of cash bond positions. For example, several traders told us about using an EFP to establish a long cash-short futures position, analogous to the speculators in our model in the previous section. In such cases, the futures position would often be unwound the same day (see also footnote 1 above).

Figures 1 and 2 show the monthly time trend of EFP volume, EFP lot sizes, and overall futures volume during 1992 to 1996. Figure 1 depicts EFPs in 3 year Commonwealth Bond futures; Figure 2 is for 10 year Commonwealth Bond futures. Table 1 provides summary statistics on the same trading activity for each year. EFP activity prior to mid 1991 (not included in our data) was very light, according to exchange sources.

EFP activity in both contracts grew at a high rate during 1992 and 1993 after Tullet and Tokyo Australia Pty Ltd, one of the nation's largest bond brokers, began sponsoring a screen-based trading facility for EFPs. For 10 year bond EFPs, trading activity continued to increase through 1994. The time trend of trading activity for both bond futures was flat or even declining in 1995 and 1996. Overall, average monthly volume of EFPs in 3 year bonds (10 year bonds) grew from about 45,000 (36,000) contracts in 1992 to 136,000 (64,000) contracts in 1996. A contract represents A\$100,000 face value of bonds.

The overall growth in EFP trading activity was due to increases in both the number of EFP trades and their average sizes. For 3 year bond EFPs, the monthly average number of trades grew from 355 in 1992 to a peak of 980 in 1994, settling to 552 in 1996. Mean monthly trade sizes rose from 132 contracts in 1992 to 255 contracts in 1996. For 10 year bond EFPs, the monthly average number of trades grew from 470 in 1992 to a peak of 754 in 1994, settling to 432 in 1996. Mean monthly trade sizes rose from 78 contracts in 1992 to 165 contracts in 1996.

Trade sizes represent one source of controversy over EFPs during this period. EFPs had traditionally been conceived as a means of effecting delivery in large-trade situations. However, in these markets, EFPs were increasingly used for standard-sized trades. A round lot trade in bonds during this period was A\$10,0000, corresponding to 100 futures contracts. In 10 year bonds, the mean EFP trade size was about 100 contracts over most of the sample period. In 3 year bonds, trades were only somewhat larger—up to around two round lots.

Sydney cash bond traders have suggested to us that EFPs offer advantages in terms of enhanced liquidity, lower transactions costs, greater flexibility in timing and more anonymity relative to straight cash market trades. Additionally, hedgers with small positions might find EFPs to be a more convenient way to close out the hedge. Conversely, traders who want to establish a small hedged position might be attracted by the exchange's regulatory or clearing

systems. The fact that many relatively small trades are done via EFPs suggests that these sorts of benefits may be very prominent in the Sydney markets.

Table 1 shows that futures volume was also increasing during much of this period, though not as fast as EFP volume. Therefore, EFP volume as a percentage of futures volume was generally increasing. EFP volume comprised 10.5 percent of 3 year bond futures volume (11.3 percent of 10 year bond futures volume) in 1992, rising to 24.1 percent (18.3 percent) in 1996.

#### 4.2 The relation of EFP trading and market volatility.

Both the model presented above and discussions with Sydney traders suggest that the avoidance of transactional risk or price slippage is an important motivation for EFP trading. Slippage is apt to be a greater concern when prices are especially volatile, so EFP trading is predicted to be most intense when volatility is high. Additionally, high volatility and slippage both correspond to high on-exchange hedging costs.<sup>11</sup> This section investigates the empirical relationship between EFP volume and volatility for SFE 3 year and 10 year bond futures, and establishes the existence of a positive relationship between EFP trading and market volatility. The purpose of this analysis is only to establish association; the issue of causality is studied in the next section.

Table 2 presents the results of regressions of daily EFP volume for each contract on daily volatility. Volatility is measured as the standard deviation of 15 minute returns in the nearest-to-expiry contract for each trading day. As before, the data cover 1992-1996.

Specification 1 of Panel A (Panel B) shows a highly-statistically-significant and positive regression coefficient on volatility equal to 0.295 (0.438). Clearly, EFP trading is most prominent on volatile days. Because of the findings above that EFP volume demonstrates a marked trend over time, the table also presents the results of estimating Specifications 2 and 3, which allow for linear and non-linear time trends, respectively. The same conclusion can be drawn from these results: EFP trading is heaviest on volatile days, consistent with the notion that EFPs are a way to avoid transactional risk.

---

<sup>11</sup> Our model suggests that the volatility of the hedged position—basis risk—is an important motivation for EFP trading. In our empirical work, we use the futures volatility as a proxy because it should be highly correlated with basis volatility, which is difficult to observe due to the paucity of intraday cash market data. As a more general justification for our use of futures volatility, note that slippage in spread trades is mainly due to the futures market, where orders must be submitted without knowledge of the exact trade price. In the cash market, negotiated prices are more common.

### 4.3 The effects of EFP trading on SFE bond futures markets

#### 4.3.1 Empirical approach and method

The analysis in the previous section does not demonstrate causality in any sense. We interpret the regressions to indicate that volatility leads to heavy EFP trading, but it is also possible that EFP trading leads to volatility. This section presents an analysis that focuses directly on the sequence of statistical causality, to better understand the effects of EFPs on the futures market and of the futures market on EFPs. The results bear directly on concerns about the effects of EFP trading on overall market quality.

To analyze the possible effects of EFP trading on the underlying futures markets, it might seem attractive to examine the market before and after EFP trading became prevalent. However, this approach leads to inconclusive results. Not only did EFP trading grow over a long time, allowing for many other influences, but the SFE bond futures markets grew dramatically at the same time. In the three years prior to 1992, average monthly volume in the 3 year bond contract was about 128,000 contracts, whereas in the three years following 1992 the analogous figure is about 586,000 contracts. For the 10 year bond contract, the comparative figures are 172,000 versus 378,000 contracts. Even if fragmentation had been drawing some volume from the futures market to the EFP market, so much additional volume was drawn to the market overall that the effect is swamped. Volatility and effective bid-ask spreads also fell over the time period (details not reported here; available on request).

Since we cannot use a before-and-after approach in searching for possibly-detrimental effects of EFP trading, we look for effects during the period when EFP trading is prevalent, January 1993 to April 1996. Our approach is to examine the time series of volatility, bid-ask spreads, on-exchange volume, and EFP volume over this period. We examine the time-series relations between the intensity of EFP trading and volatility (as a proxy for on-exchange slippage and hedging costs), bid-ask spreads (as a proxy for liquidity), and on-exchange volume. Recall that the model presented above indicates that a fragmented market can be detrimental to hedging costs, liquidity, and on-exchange volume. In contrast, the model also shows that these detrimental effects can be offset in an EFP market without transactional risk. The point of our investigation is to determine if the SFE EFP market looks more like a detrimental fragmented market, or a market with neutral-to-beneficial EFPs.

Daily observations on futures volume, EFP volume, volatility, and the bid-ask spread are required for this investigation. The first two are directly available in data from the SFE. As before, volatility for each day is measured as the sample standard deviation of returns at 15 minute intervals.

Since the bid-ask spread is not directly observable in the transaction-price record available from the exchange, it must be estimated. A variety of estimators have been proposed for futures data like ours, including Roll (1983) and Smith and Whaley (1994). The Roll estimator is known to be highly inefficient, and the Smith and Whaley estimator requires a normality assumption, which we have found to be rejected for our data (details not reported; normality is commonly rejected for high-frequency futures data, so this is not surprising). Therefore we use a new method, the details of which are given in the Appendix. The basic logic of the method is that the price changes over a longer interval (say, 5 minutes) impound both the bid-ask bounce effect of the spread and the effect of fundamental volatility, while price changes over a shorter interval (say, 1 minute) impound proportionately less influence of fundamental volatility but the same spread effect. Thus, comparing the measured variances, one can extract an estimate of the spread. This estimate, computed for each day with data from the nearest-to-expiry futures contract, is the measure of the spread we use.

For each futures contract, we estimate a vector autoregression (VAR) for the joint daily time series of futures volume, EFP volume, volatility, and spread. VAR is a way of estimating dynamic relationships among jointly-endogenous variables such as these, without imposing strong a priori restrictions (see, for example, Sims (1980)). We include dummy variables for the expiry months of March, June, September, and December in case futures or EFP volume changes deterministically near expiration. Our VAR includes five lags of every jointly-endogenous variables, on the assumption that observations more than a week apart (five trading days) are unlikely to directly influence each other.

From the results of the VAR, we compute a set of impulse-response patterns for each of the four jointly-endogenous variables (futures volume, EFP volume, volatility, and spread). These impulse-response patterns show how each particular variable responds over time to an exogenous shock in each other variable.<sup>12</sup> Impulse-response patterns thus provide a means of determining whether one variable statistically causes another variable. In this context, we are particularly interested in whether increases in EFP trading cause decreases in futures volume, increases in volatility, or increases in spread.

### 4.3.2 Results

The results of this investigation for the 3 year bond market are shown in Figure 3. Results for the 10 year bond market are shown in Figure 4. The figures contain a panel for every possible combination of causing and caused variables. The central plotline in each panel

---

<sup>12</sup> The results in Figures 3 and 4 are generated by simulating one-standard-error shocks, respectively, to each variable of interest. In each case, “standard error” refers to the root of the mean square residual from the appropriate VAR equation.

shows the best estimate of the impulse-response in the “Response of” variable to a one standard-deviation shock in the “To increase in” variable. The upper and lower plotlines in each panel show two-standard-error bounds on the response, computed by a standard Monte Carlo simulation.

Figure 3, for the 3 year bond market, shows that many dynamic effects cause other effects in this market. For example, the third panel in the third row, labeled “Response of Volume to Increase in Volume” shows how volume shock are propagated across days.<sup>13</sup> A one-standard-deviation shock to volume, which is about 12,000 contracts on a given day, results in increased volume on most of the next 11 days (i.e., with two-standard-error bounds above zero; the bounds touch zero on days 3, 8, and from day 12 onward).

The fourth row of Figure 3 is most central to our investigation. This row depicts the response of each jointly endogenous variable to a one-standard-deviation shock in EFP volume (of about 2500 contracts). Note that a shock in EFP volume leads to more EFP volume on subsequent days—the fourth panel shows that the “Response of EFP Volume to Increase in EFP Volume” is reliably positive for each of the following 8 days.

In contrast, no other variables (spread, volatility, or volume) show any significant contemporaneous or persistent response to the “Increase in EFP Volume.” Thus, the investigation shows no evidence of detrimental effects of increased EFP trading on aspects of the on-exchange futures market.<sup>14</sup>

Notably, there is a significant persistent response of EFP volume to an increase in futures volume. This suggests that the EFP market may be an outlet for excess volume during periods of stress in the on-exchange market.

Figure 4, showing impulse-response diagrams for the 3 year bond market, shows almost exactly the same effects as for the 10 year bond market. However, the tendency for EFP

---

<sup>13</sup> The 3 year bond (only) shows a puzzling regularity in Figure 3. In that figure, a large effect at lag 5 is present in all cases that have a large day-0 effect. To check that this is not mechanically induced by our choice of 5 lags in the VAR, we have re-computed the results with various other lag structures. There are no systematic changes in the results. For example, with 6 lags in the VAR, all but one of the original lag 5 effects remain essentially unchanged. These results, along with the fact that there is no corresponding regularity for the 10 year bond, lead us to believe that our lag structure is not the cause.

<sup>14</sup> A one-standard-error shock to daily EFP volume is only about 1/6 of an average day’s futures volume. In case this shock is too small to imply an economic response, we have also computed impulse-response functions for larger shocks. No additional statistically significant effects are detected. Our large-shock results should be interpreted cautiously, however; if dramatic increases in EFP trading were encountered, the coefficients and distributional characteristics of the VAR would perhaps change.

volume to lead to subsequent EFP trading is more pronounced in the three-year bond market.<sup>15</sup> Similar findings for both markets adds weight to our conclusions.<sup>16</sup>

## 5. Conclusion

In recent years, alternatives to trading on a centralized exchange have proliferated. In the financial futures markets, the exchange of futures for physicals (EFPs) has become increasingly popular over the past decade. EFPs are a private means of exchanging futures contracts for cash-market equivalents.

A major attraction of EFPs is that, unlike on-exchange trading, they allow the price and terms of a possibly-complex transaction to be fixed in advance. Traders often mention the avoidance of transactional risk (slippage) as one of the major attractions. Although clearly very useful to some traders, EFP trading has raised concerns about whether the availability of this alternative trading mechanism might be detrimental to the volume and the liquidity of the futures markets.

The model in this paper points out that, while fragmenting the futures market can lead to clear harm to the primary futures market, a tightly-linked EFP market can, by allowing for the avoidance of transactional risk, offset that harm. Given a well-known opportunity to trade without transactional risk, traders are less inclined to back away from taking positions on the exchange.

It is an empirical issue whether EFPs merely fragment the market, or, alternatively, provide some offsetting benefits. Using a vector autoregression approach, we examine two government bond contract markets at the Sydney Futures Exchange. Since the futures market is the only centralized public market related to the government bonds we study (trading in physical bonds is over the counter), the links represented by EFP trading may be all the more significant. We find no negative effects of EFP trading. Once begun, EFP volume does persist over time, but there is very little evidence of accompanying or subsequent detrimental effects on futures market volume, volatility, or liquidity.

---

<sup>15</sup> The 10 year bond market shows evidence of the well-known tendency for volume to cause volatility. This tendency is not evident in the 3 year bond market.

<sup>16</sup> Exchange regulations require that EFPs be reported within  $2\frac{1}{2}$  hours. Given that EFPs are negotiated away from the Exchange, such requirements are difficult to enforce and some late-in-the-day EFPs could have a next-day time stamp. To investigate whether this might have affected our results, we have recomputed the VAR and its impulse-response representation using data for rolling 2-day periods. The two-standard-error bounds widen, as should be expected with fewer observations, but essentially no other changes are observed.

We observe that the relationship between the on-exchange futures market and the EFP alternative is in some respects complementary. EFPs reference the futures price, thus making use of the price discovery on the exchange. The services of the futures clearinghouse are essential to EFP counterparties, and the clearinghouse guarantee cannot be strong if the exchange is not healthy. Also, EFP trading often results in open futures positions that will later be traded out, so EFP traders have a material interest in the liquidity of the overall futures market. Finally, if the availability of EFP trading makes futures a more attractive risk management tool, then more users are attracted to the marketplace.

From a public policy perspective, the results suggest that the availability of alternative markets might not presage the demise of centralized markets. Outcomes in the centralized market need not be harmed by the availability of an alternative trading mechanism. On the regulatory front, our results suggest that exchanges should push for tight linkages with alternative markets like those for EFPs. For example, real time reporting of EFP trading information helps insure against detrimental effects and enhances the possibility of more volume being attracted to a more useful futures market.

**APPENDIX: MEASURING THE BID-ASK SPREAD IN FUTURES MARKETS**

In futures markets, the bid and ask prices available to arriving traders are not generally recorded unless no trades are occurring. Thus, the bid and ask prices at which transactions actually occur are not available for research purposes. What is recorded are the prices at which trading takes place. This record is known as “PIT” data in Sydney markets.<sup>17</sup> Analysis of the bid-ask spread, then, must be accomplished by way of inference from the record of traded prices.

A variety of measures have been proposed to infer spreads from the time series of trade prices. Roll (1983) suggests a measure based on the serial covariance of returns. Assuming the direction of successive trades is random, the serial covariance registers the magnitude of the bid-ask bounce in the data, allowing the spread to be inferred. Roll’s measure, though conceptually simple, has a practical disadvantage: it produces no result when the serial covariance is positive (i.e., when there is no bid-ask bounce). The serial covariance can be positive when the direction of successive trades is autocorrelated or whenever prices are trending.

Alternatively, the mean absolute value of price changes has been suggested as a measure of the spread, since that statistic measures the extent of the price move across the bid-ask spread. The disadvantage of this measure is that the absolute size of price moves is a function of fundamental volatility as well as of the spread. When futures prices are volatile, the mean absolute price change will be large irrespective of the spread.

Smith and Whaley (1994) suggest a procedure that simultaneously estimates, from the time series of prices, the spread and the fundamental volatility. Their method assumes that fundamental price changes are normally distributed. This adds sufficient structure to the problem so that they can solve for spread and volatility from the system of equations that theoretically describes the mean and mean square of absolute price changes.

This paper uses a new method of estimating the spread that improves on these previous approaches. It bears some similarities to Smith and Whaley’s approach but does not require the assumption of normality. Additionally, the method developed here uses the full information content of the available data to estimate the spread, rather than attempting to additionally estimate the fundamental volatility.

The basic idea is that tick-to-tick price changes impound the influences of both the spread and of fundamental volatility. Price changes over longer periods, say five minutes, impound the same spread effect; however, because more time and trades have passed, the potential for fundamentally driven price changes is increased. The extent of the increase is indexed by the extent of trading in the interval, based on the well-established idea that extensive trading coincides with intense information arrival. By comparing the amount of trading and the extent of price changes over five-minute periods with those in tick-to-tick data, the

---

<sup>17</sup> The Sydney Futures Exchange also track prices and volume via records used for clearing purposes. These data are not useful for inferring liquidity, as the time stamps are potentially inaccurate and the records do not indicate which party initiated the trade.

influence of fundamental volatility can be calculated and purged. What remains is an estimate of the spread.

Formally, assume (like Roll 1983; Smith and Whaley 1994) that the  $t$ -th tick-to-tick price change is described by

$$\Delta P_t = sI_t + e_t , \quad (\text{A1})$$

where  $s$  is the bid-ask spread;  $I_t$  is an indicator variable equal to -1.0 for bid-to-ask price changes, equal to 1.0 for ask-to-bid price changes, and equal to 0 otherwise; and  $e_t$  is a mean-zero fundamental price change with variance  $\sigma^2$ . The direction of successive trades is assumed to be independent, so the probabilities that  $I_t$  equals -1.0 or 1.0 are one-quarter each, and the probability that  $I_t$  equals 0 is one-half.

It follows from equation (A1) that the variance of price changes over  $k$  minutes is

$$\text{Var}(\Delta P_k) = T_k \sigma^2 + \frac{1}{2} s^2 ,$$

where  $T_k$  is the number of trades over the  $k$ -minute interval.

To estimate the spread, choose two separate intervals over which to measure price change variance, and then solve for the spread, using equation (A2). The intervals used in this paper are  $k = 1$  minute and  $k = 5$  minutes. The resulting expression for the spread is

$$s = \sqrt{\frac{2(T_5 \text{Var}(\Delta P_1) - T_1 \text{Var}(\Delta P_5))}{T_5 - T_1}} . \quad (\text{A2})$$

This estimator does not require the assumption of normal fundamental price changes. However, if fundamental price changes are normally distributed, then  $s^2$  is distributed as a  $\chi^2$  random variable, which allows for the use of standard statistical tests.

## REFERENCES


- Biais, B. 1993. "Price Formation and Equilibrium Liquidity in Fragmented and Centralized Markets." *Journal of Finance* 48(1): 157-185.
- Greenwald, B., and J. Stein. 1991. "Transactional Risk, Market Crashes, and the Role of Circuit Breakers." *Journal of Business* 64(4): 443-62.
- Hendershott, T., and H. Mendelson.. 2000. "Crossing Networks and Dealer Markets: Competition and Performance." *Journal of Finance* 55 (5): 2071-2116.
- Hogan, R. 2000. "New T-corp Benchmark." *The Australian Financial Review*. April 11: 51.
- Madhavan, A. 2000. Market Microstructure: A Survey. *Journal of Financial Markets* 3(3): 205-258.
- O'Hara, M. 1995. *Market Microstructure Theory*. Blackwell Publishers.
- Pagano, M. 1992. "Trading Volume and Asset Liquidity." *Quarterly Journal of Economics* 104:255-74.
- Redheffer, N. 1990. "Exchange for Physicals in the Chicago Board of Trade US Treasury Bond Futures Contract." CFTC Memorandum, July 20.
- Roll, R. 1983. "A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market." *Journal of Finance* 39(4): 1127-39.
- Smith, T., and R. Whaley. 1994. "Estimating the Effective Bid/Ask Spread from Time and Sales Data." *Journal of Futures Markets* 14(4  CD Player.Ink ): 437-455.

Table 1. Yearly averages of monthly trading volume, number of trades, and mean trade size in the 3 and 10 year Commonwealth Bond futures market, 1992-1996. Except for percentages, all items are measured as number of contracts. Data for December 1992 and December 1993 are missing. 1996 entries are based on data for the first quarter only.

## Panel A. 3 Year Commonwealth Bond Futures

Year	Monthly average EFP volume	Number of EFP trades per month	Average EFP trade size	Monthly futures volume	EFP volume as percentage of futures
1992	44660	355	132	425483	10.5%
1993	86241	782	110	524967	16.4%
1994	134927	980	138	674700	20.0%
1995	146899	867	173	598068	24.6%
1996	135779	552	255	563304	24.1%
All Years	106219 (45429)	740 (283)	145 (41)	558877 (162169)	19.0

## Panel B. 10 Year Commonwealth Bond Futures

Year	Monthly EFP volume	Number of EFP trades per month	Average EFP trade size	Monthly futures volume	EFP volume as percentage of futures
1992	36169	470	78	321087	11.3%
1993	58821	742	79	355313	16.6%
1994	74266	754	98	448009	16.6%
1995	78644	694	114	371192	21.2%
1996	64026	432	165	349958	18.3%
All Years	60131 (22274)	620 (219)	97 (26)	365540 (98202)	16.5%

Table 2. Regressions of Daily EFP Volume on Daily Volatility in the 3 and 10 year Commonwealth Bond futures market, 1992-1996. Daily EFP volume is measured as number of contracts, and daily volatility is measured as the standard deviation of returns over 15-minute intraday periods. Data for December 1992 and December 1993 are missing. t-statistics are in parentheses.

	Intercept	Coefficient on Daily Volatility	Coefficient on Time Trend $\times 1000$	Coefficient on Squared Time Trend $\times 10,000$	R-square (Adjusted R- square)
Panel A. 3 Year Commonwealth Bond Futures (804 observations)					
Specification 1	7.460 (40.74)	0.295 (5.86)			0.041 (0.040)
Specification 2	6.905 (38.58)	0.345 (7.29)	0.927 (10.72)		0.161 (0.159)
Specification 3	6.796 (38.75)	0.296 (6.34)	0.338 (9.02)	-0.027 (-6.49)	0.203 (0.200)
	Intercept	Coefficient on Daily Volatility	Coefficient on Time Trend $\times 1000$	Coefficient on Squared Time Trend $\times 10,000$	R-square (Adjusted R- square)
Panel B. 10 Year Commonwealth Bond Futures (805 observations)					
Specification 1	5.970 (24.50)	0.438 (8.14)			0.076 (0.075)
Specification 2	5.575 (23.13)	0.456 (8.51)	0.327 (3.67)		0.091 (0.089)
Specification 3	5.683 (22.97)	0.433 (8.10)	1.65 (4.67)	-0.165 (-3.87)	0.108 (0.105)

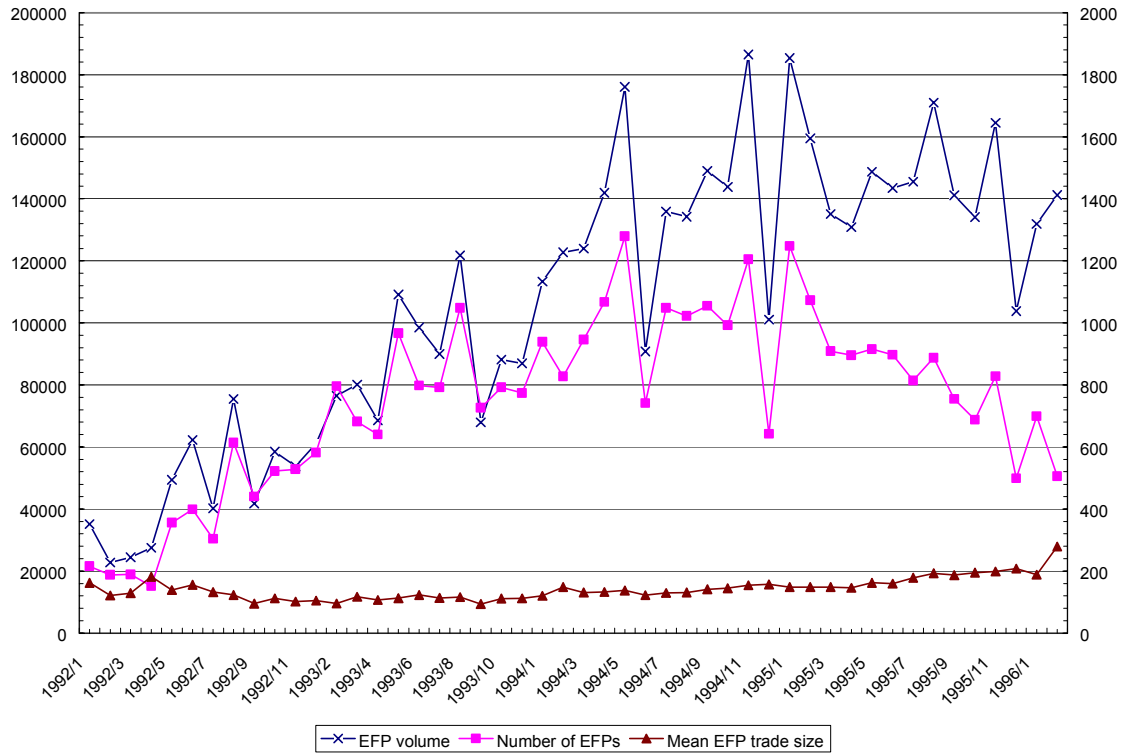


Figure 1. Monthly trading volume, number of trades, and mean trade size in the 3 year Commonwealth Bond futures market, 1992-1996. Volume, in number of contracts, is measured on the left axis. Number of trades and mean trade size, in number of contracts, are measured on the right axis. Data for December 1992 and December 1993 are missing.

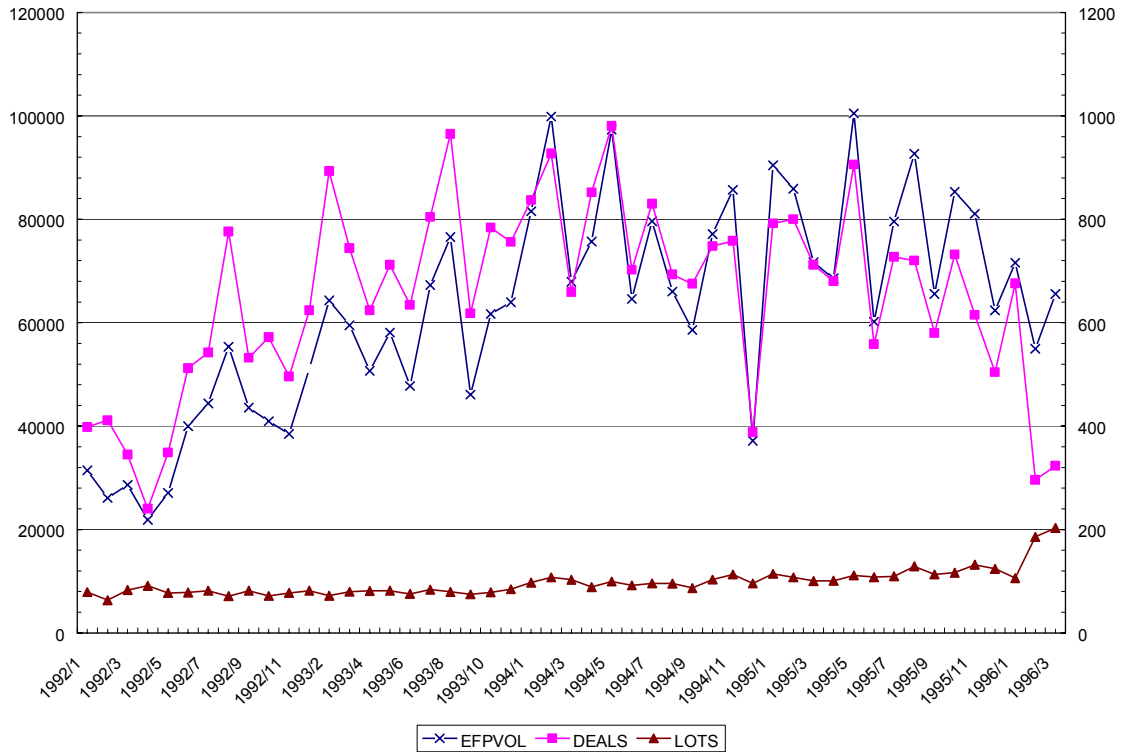


Figure 2. Monthly trading volume, number of trades, and mean trade size in the 10 year Commonwealth Bond futures market, 1992-1996. Volume, in number of contracts, is measured on the left axis. Number of trades and mean trade size, in number of contracts, are measured on the right axis. Data for December 1992 and December 1993 are missing.

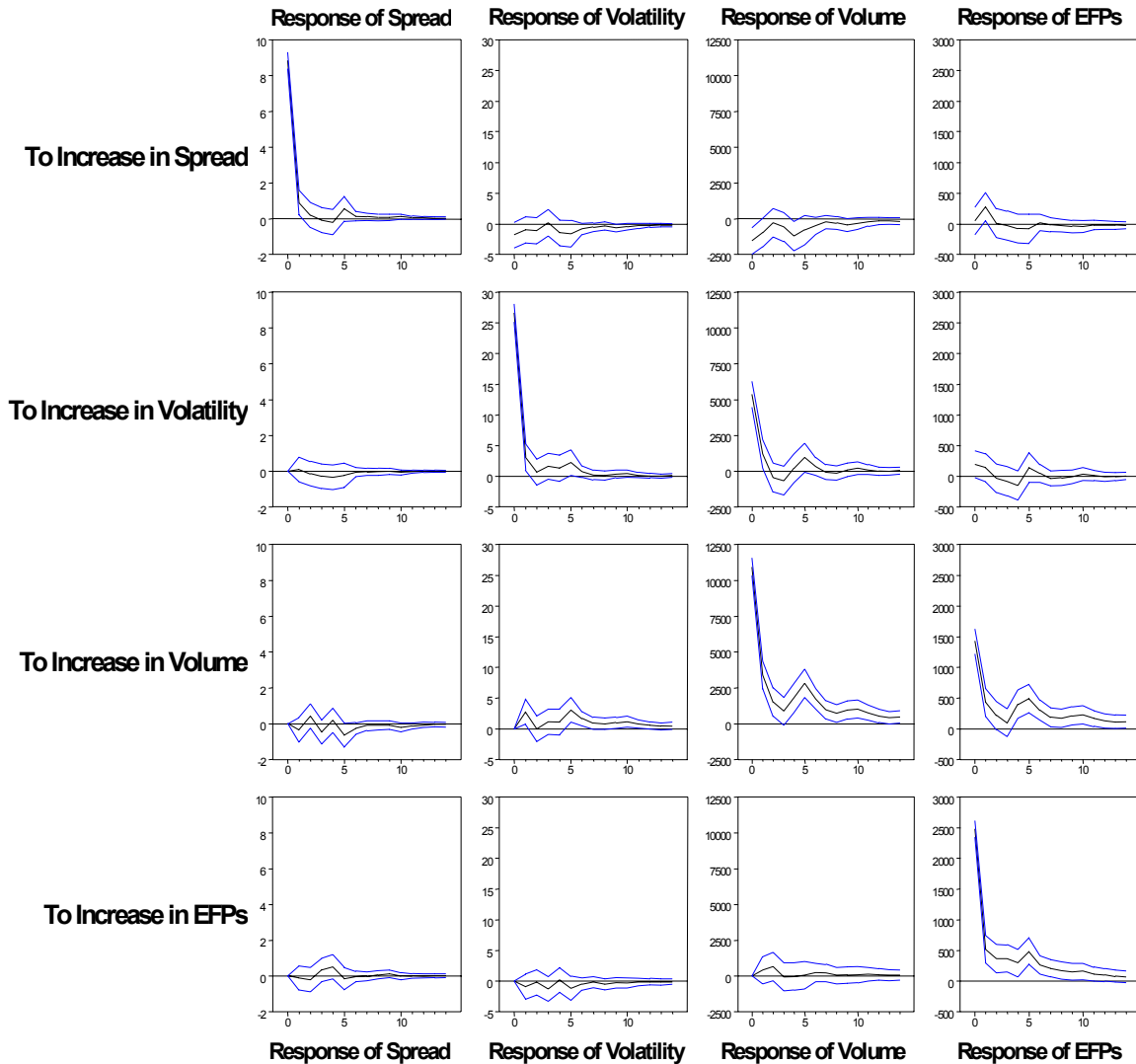


Figure 3. Impulse-Responses for the SFE 3 year Commonwealth Bond Futures market. Each impulse-response diagram depicts the response in one variable to a one-standard-error exogenous shock in some variable. The impulse-response diagrams are computed from a fifth-order vector autoregression (VAR) augmented by expiration-month dummy variables. Daily data on spread, volatility, futures volume, and EFP volume from January 1993 to April 1996 is used to compute the VAR. The horizontal axis in each panel is number of days from the day of the exogenous shock. The vertical axis is, respectively, Australian dollars for the spread and the volatility, and number of contracts for volume and EFP volume. The central curve in each plot shows the best estimate of the impulse-response path. The upper and lower curves are two-standard-error bounds, computed by standard Monte Carlo methods.

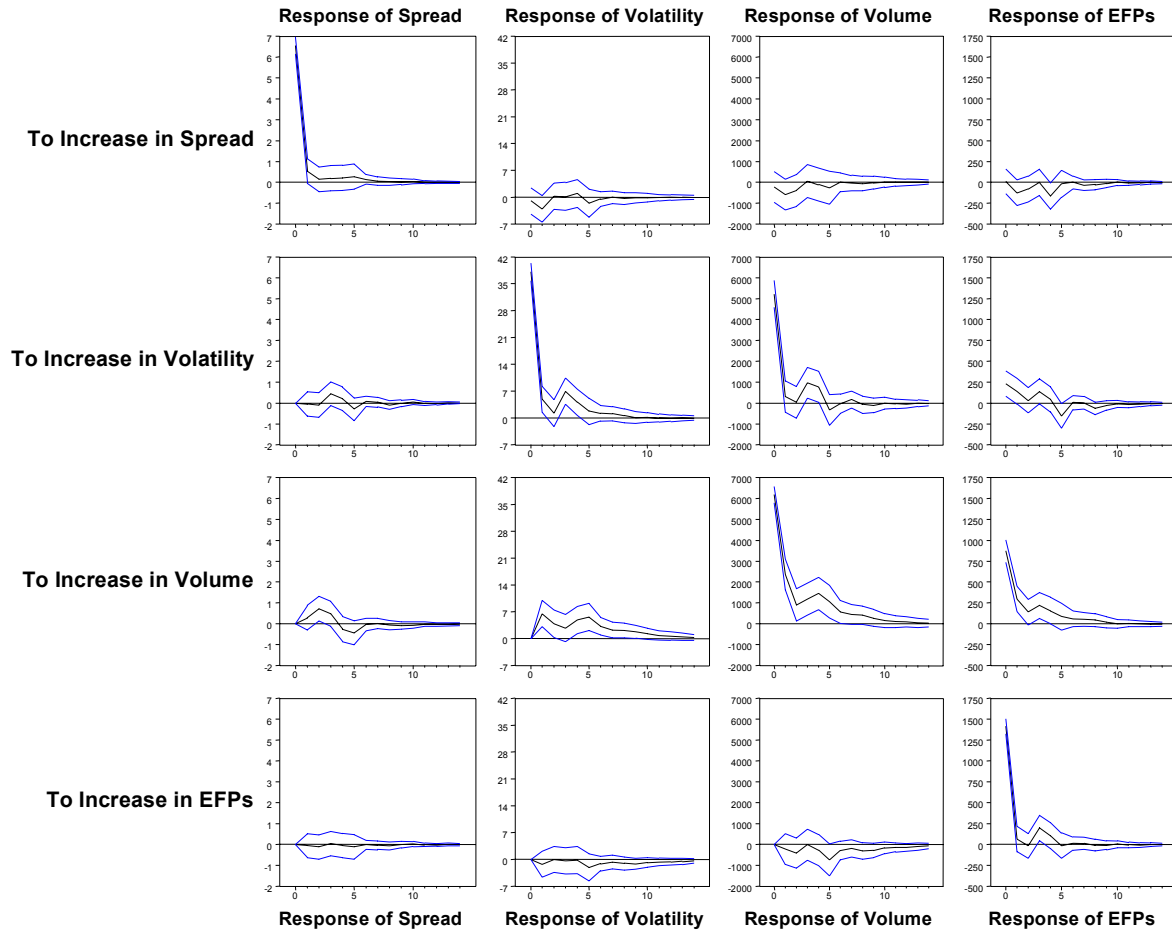


Figure 4. Impulse-Responses for the SFE 10 year Commonwealth Bond Futures market. Each impulse-response diagram depicts the response in one variable to a one-standard-error exogenous shock in some variable. The impulse-response diagrams are computed from a fifth-order vector autoregression (VAR) augmented by expiration-month dummy variables. Daily data on spread, volatility, futures volume, and EFP volume from January 1993 to April 1996 is used to compute the VAR. The horizontal axis in each panel is number of days from the day of the exogenous shock. The vertical axis is, respectively, Australian dollars for the spread and the volatility, and number of contracts for volume and EFP volume. The central curve in each plot shows the best estimate of the impulse-response path. The upper and lower curves are two-standard-error bounds, computed by standard Monte Carlo methods.