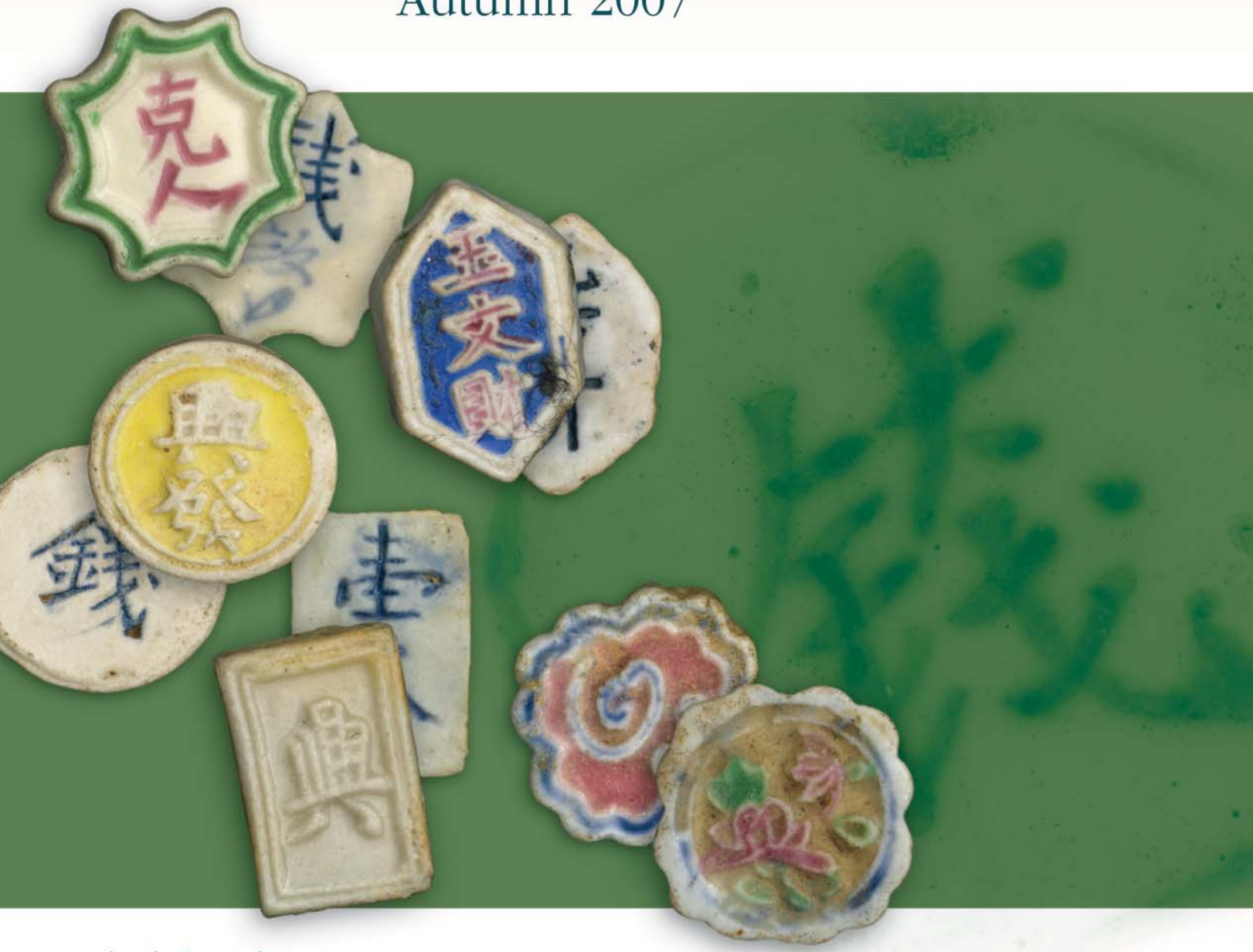
BANK OF CANADA
BANQUE DU CANADA

Bank of Canada Review

Autumn 2007



www.bankofcanada.ca

Bank of Canada Review

Autumn 2007

Articles

Cover

Siamese Gaming Tokens

Paul Berry, Chief Curator, Currency Museum

During the late eighteenth and early nineteenth centuries, there was a shortage of small-value coins in Siam (now Thailand), especially in the immediate area of Bangkok, the capital city. To meet their needs for small change, people began to use tokens from the country's numerous gambling houses for their everyday transactions. Originally used for games like *fan- tan*, the tokens were made of glass or metal until the early nineteenth century, when cheaper glazed porcelain varieties were introduced from China.

Cast in moulds, the tokens are simple in design. On one side appears the name of the issuing house, or *hong*, or an expression of good fortune, such as "prosperity," "success," or "abundance." Occasionally, this text is supplemented by an image of an animal, flower, or another object. An indication of the token's value, painted in glaze, often appears on the reverse of the piece.

Inscriptions were usually rendered in Chinese, the language of the ethnic group that owned most of the gambling houses, but on occasion the value of the token was also shown in Siamese. The tokens were denominated in the values of the local currency, which could be converted to values in cowrie shells or silver.

Ranging in size from 10 mm to 30 mm in diameter, Siamese gaming tokens are a visual delight. The tokens may be plain white, white with a blue-glaze inscription, or coated with a multitude of coloured glazes. They also come in a variety of shapes. Although most frequently round or polygonal, tokens can have scalloped edges or be in the shape of diamonds, ovals, and other irregular forms such as bottles, scrolls, leaves, and feathers.

There are thousands of known designs because the owners of the gambling houses regularly issued new tokens, both to obstruct counterfeiters and to raise profits. Designs were sometimes changed overnight, creating substantial profits for the gambling houses, since they were not obliged to redeem obsolete issues. With the release of regular government copper coinage in 1874, further issues of the gaming tokens were prohibited. Documentary evidence suggests, however, that the tokens continued to be used for decades.

The tokens pictured here are part of the National Currency Collection, Bank of Canada.
Photography by Gord Carter, Ottawa.

The *Bank of Canada Review* is published quarterly, in print, and on the Bank's website (www.bankofcanada.ca). *Banking and Financial Statistics* is published monthly. Subscriptions are available to both publications.

***Bank of Canada Review* (quarterly)**

Delivery in Canada	CAN $25
Delivery to the United States	CAN $25
Delivery to all other countries, regular mail	CAN $50

Canadian government and public libraries and libraries of Canadian and foreign educational institutions may subscribe at one-half the regular price. Single copies of the quarterly *Review* are $7.50. Single copies of the *Banking and Financial Statistics* are $5.00.
Reprints of articles are available at $2.00 per copy.

Subscriptions or copies of Bank of Canada documents may be obtained from Publications Distribution, Communications Department, Bank of Canada, Ottawa, Ontario, Canada K1A 0G9; telephone 613 782-8248, toll free in North America: 1 877 782-8248; email address: publications@bankofcanada.ca. Remittances in Canadian dollars should be made payable to the Bank of Canada. Canadian orders must add 6 per cent GST as well as PST where applicable.

Inquiries related to interest rates or exchange rates should be directed to 613 782-7506.

The Global Foreign Exchange Market: Growth and Transformation

William Barker, Financial Markets Department

- *The foreign exchange market is in a period of remarkable transformation that is changing who is trading, why, and how.*
- *Not only are trading volumes in the foreign exchange market expanding rapidly, the primary sources of this market growth are helping to define the profound structural transformation taking place.*
- *Deal flow in the foreign exchange market is increasingly transacted electronically, using automated computerized trading routines, and by a much wider array of market participants.*
- *These changes reflect innovative developments in electronic trading technology and institutional trading arrangements that are shifting the balance of market participation between bank and non-bank accounts, large and small market participants, and domestic and global players.*
- *As a result of this ongoing structural evolution, the foreign exchange market is arguably becoming more liquid and operationally efficient.*

The foreign exchange market is known to be the largest financial market in the world, as measured by daily turnover. The most recent BIS Triennial Survey (BIS 2007) estimates that the total turnover in global foreign exchange markets is US$3.2 trillion a day—more than 6 times larger than trading in U.S. Treasury bonds and 30 times larger than trading on the New York Stock Exchange (SIFMA 2007; NYSEData.com 2007). What may be less apparent is how quickly this market has grown over the past few years and why it is growing so quickly. By most estimates, the trading volumes in the foreign exchange market are continuing to grow rapidly. The Tower Group, for example, recently estimated that daily global trading volumes would likely reach US$5 trillion by 2010 (*Profit & Loss Magazine* 2007). Should this happen, foreign exchange trading volumes will have more than tripled in this decade (Chart 1).

As the market expands, it is undergoing a remarkable transformation that is changing who is trading foreign exchange and how they are doing it. In turn, these changes are accelerating market growth. An increasing proportion of overall foreign exchange trading volume is being transacted on electronic trading platforms, both in the interbank market as well as between banks and clients; by large global investment dealers and non-bank market participants; and by computer-driven algorithmic trading strategies. Together, these closely related and mutually reinforcing elements are defining a new paradigm for the foreign exchange market and, indeed, for global financial markets in general.

Chart 1

Global Average Daily Trading Volume in Foreign Exchange

Sources: Price (2007)
BIS, April 2007

This article describes how these factors are changing the structure of the foreign exchange market, from a model that prevailed as recently as the mid-1990s to a new model that is still evolving. It is important to emphasize that the foreign exchange market is still in a period of transition: while the market as a whole, and particularly its most rapidly growing sectors, may be moving towards a new trading model, different market participants are nonetheless arrayed at different points along a spectrum of change. As a result, the foreign exchange market is currently a mixture of old and new elements.

We begin describing this evolving mixture by examining the factors behind the strong growth in the foreign exchange market: changing technology, the opening of access to the market to a broader range of participants, and the automation of trading functions. This is followed by a review of how these factors have impacted market liquidity and operational efficiency.

Innovations in Electronic Dealing Technology

Through the mid-1990s, the foreign exchange market was primarily reliant on phone-based technology. A client needing to deal in foreign exchange would phone a bank with whom it had a line of credit and ask for a two-sided price, i.e., a bid and offer on the specified amount of foreign exchange to be transacted.[1] Banks would quote prices for their clients on demand, serving as market-makers. The market-maker bank would have typically transferred, or laid-off, the risk created by the deal in the interbank market by phoning other banks with which it had established a dealing relationship and conducting an offsetting transaction. (These interbank dealing relationships were mutual obligations between banks to quote each other two-sided prices on demand for wholesale amounts of foreign exchange—typically US$5 million or larger.) This phone-based network of direct relationships between banks was the principal component of the interbank market, the central source of liquidity in the foreign exchange market.[2] Frequently, banks' participation in these interbank dealing relationships was motivated solely by price discovery. Because the wholesale price apparent to a dealer consisted only of the two-sided quotes provided on demand by other banks (and even then, only for the duration of the phone call) and because they were faced with a constantly changing foreign exchange rate, banks were forced to make frequent calls to each other throughout the business day to learn the current wholesale price. Banks would typically "pay away the spread" (the difference between bid and offer quotes) on price-discovery transactions as a necessary cost of doing business.

During the past decade, these interbank dealing arrangements began to shift to electronic protocols. Reuters Dealing and EBS (Electronic Broking Services) both introduced electronic interbank trading platforms in the early 1990s.[3] Although uptake of electronic broking was relatively slow at first, by the late 1990s these platforms came to dominate interbank trading flows.[4] By most estimates, their combined market share now accounts for about 90 per cent of interbank trading in most major currency pairs; voice broking accounts for

1. The protocol was usually to ask for a two-sided price, rather than to indicate whether the transaction would be a purchase or sale, so that the dealer would not "shade" the price against the client.

2. Alternatively, banks could use voice brokers (so called because they used "squawk boxes" that gave a live audio feed on available prices to their client banks) as intermediaries in deals with other banks. Voice brokers would search for dealing interest among their client banks for transactions at a given price or amount, collecting a proportional fee every time a deal was completed. They acted solely as agents, and the prices they quoted were valid only in the size and for the amount of time determined by the banks acting as principals.

3. The EBS platform was introduced in September 1993, and Reuters Dealing 2000–2 in 1992. The 2000–2 version has live streaming prices, but an earlier version using electronic messaging for trading had been in place since the 1980s. In 2007, EBS was bought by ICAP, a large interbank fixed-income broker and was renamed ICAP EBS.

4. These two interbank platforms dominate trade in different currency pairs rather than competing directly against each other. EBS, the largest, carries most of the interbank trading volume in the euro, yen, and Swiss franc; Reuters Dealing dominates trade in the pound sterling and the Canadian, Australian, and New Zealand dollars, as well as in several emerging-market currencies.

most of the rest, while direct dealing among major banks has all but disappeared.

The price-discovery process on Reuters Dealing and EBS differs from the phone-based model of direct dealing in several key aspects. First, banks participating on these platforms are not obliged to provide two-sided price quotes to other banks on demand. A bank can post a one-sided price (either a bid or an offer) and only when it chooses to. Second, the minimum deal size allowed on these portals is much smaller than the standard wholesale amounts used in the traditional direct-dealing relationships between banks. This allows any dealer with a smaller amount to transact to enter the market without the obligation of making or accepting delivery of unwanted, larger amounts. Third, and perhaps most importantly, these electronic portals provide a live price stream that aggregates all bids and offers posted on the system. This interbank price is visible at all times to all participating dealers.

The same technology that enabled electronic price delivery in the interbank market was relatively easily extended to bank-to-client (B2C) relationships as well. Single-bank portals are bank-owned trading platforms that establish an electronic communications link between the dealer and its end-user clients, supplying that dealer's price quotes and trade details electronically. Multi-bank portals are third-party platforms that connect an end-user client with price quotes from several banks simultaneously. (Examples of multi-bank portals include FX Connect and FXall; average daily trading volumes on these two platforms are shown in Chart 4, below.) The technology for both single and multi-bank portals now makes live, dealable, streaming price quotes (similar to those available on electronic interbank platforms) available to end-user accounts.

A Changing Mix of Market Participants

Technological innovation dramatically reduced trading costs and has created new opportunities, as well as new challenges, for a broad range of market participants. This has occurred in several principal ways.

First, the ability to transact in relatively small amounts on fully transparent prices on these global electronic dealing platforms has led to fundamental changes in the operation of the interbank market. Since banks no longer need to engage in costly price-discovery transactions and mutual dealing relationships with other dealers, the foreign exchange market has been opened up to much broader global participation among banks in the provision of liquidity. The electronic aggregation of a multitude of worldwide orders and transparency in pricing has also led to sharp compression in the typical interbank bid / offer spread.

Technological innovation dramatically reduced trading costs and has created new opportunities, as well as new challenges, for a broad range of market participants.

Second, heightened competition between dealers and the much greater degree of price transparency has led to interbank spread tightening being passed along to end-user accounts in the B2C market. As dealing costs for end-user clients have declined, new accounts entered the foreign exchange market, and existing market participants were able to profitably transact more. Increased trading by end-user clients was further facilitated by the cost efficiencies of using B2C dealing portals. The use of electronic foreign exchange trading by buy-side accounts has been growing steadily: it is estimated that in 2006, for the first time, more than half of all foreign exchange transactions by end-user clients were executed electronically (Greenwich Associates 2007).

Third, many of the market-making banks that previously dominated the market have been forced to re-examine their business model as dealing spreads in both the interbank and B2C markets compress. The result has been a changing mix of large and smaller banks in the interbank market. Technology is expensive. In addition to considerable start-up costs, it requires continuous upgrades. Given the tighter bid / offer spreads in both the interbank and B2C markets, most successful dealing banks have therefore implemented a low-margin, high-volume business model that amortizes these higher technology costs through continually building trading volumes. This gives a competitive advantage to those banks with the size and large global distribution networks needed to sustain ongoing technological innovation and to provide competitive, profitable price quotes. The result has been consolidation in the foreign exchange market, with the largest banks accounting for a growing percentage of the overall global trading volume. For example, in the May 2007 *Euromoney* foreign exchange poll, five

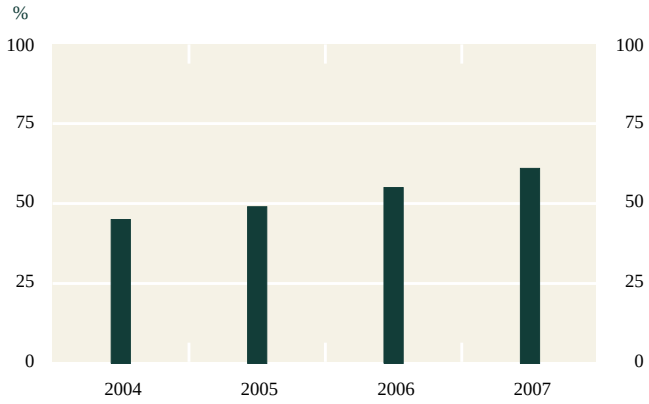
banks accounted for over 60 per cent of client trading activity.[5] (See Chart 2.) In the 2006 poll, the top five had a 54 per cent market share; a decade ago, the top five accounted for less than one-third of market volume.

Fourth, as a result of deal flow in the global foreign exchange market consolidating among the largest global banks, the role of second-tier dealers has been evolving. For smaller banks, the level of technological commitment needed to remain competitive in such a low-margin environment or to operate in all currency pairs and all time zones is no longer feasible. It often makes more economic sense for them to outsource this function to large global institutions.

For some second-tier dealers, this outsourcing has taken the form of "white-labelling" (or white-boarding), whereby the smaller bank will act as an intermediary between an end-user client buying foreign exchange and another larger bank that supplies it. Essentially, the smaller bank becomes a liquidity retailer, maintaining its single-bank B2C portal for servicing client orders but using a larger bank to provide the wholesale liquidity.[6] The smaller bank is thus able to specialize in managing the client's credit risk. The larger bank provides the liquidity and manages the market risk generated by the client's order. This institutional division of labour and specialization in areas of comparative advantage supports better pricing for the end-user client.

Chart 2

Combined Market Share of the Five Largest Banks in the Foreign Exchange Market



Source: *Euromoney Magazine* (2007)

While white-labelling allows liquidity outsourcing in the B2C market, a similar institutional dealing arrangement —prime brokerage—allows outsourcing of liquidity provision in the interbank market. Many smaller dealers do not have access to a broad range of reciprocal credit lines and thus to the most competitive interbank price quotes. Prime brokerage arrangements allow such dealers to access the interbank market by using the credit relationships of a top-tier bank. The larger bank acts as an intermediary, connecting the smaller dealers with competitive interbank market pricing and managing their credit risk, but not assuming any price risk itself. (As with white-labelling, the intermediary in prime brokerage simply passes through the pricing and price risk to the client.) Prime brokered dealing typically provides the smaller bank with better pricing than it could obtain on its own; the prime broker, in turn, earns a fee for this service.

Fifth, prime brokerage has recently created new trading opportunities for market participants outside the banking sector. Although prime brokerage originated as a specialized relationship between dealers in the interbank market, in the past couple of years these dealing relationships have been extended to a large and growing class of market participants largely outside the banking sector, the professional trading community (PTC). Hedge funds, in particular, have grown enormously in the past decade, both in numbers and in the amount of capital under management, to form the core of the PTC. The PTC also includes commodity-trading advisers that manage exchange-trade futures accounts, as well as currency-overlay managers that actively manage foreign exchange exposures in investment portfolios. Significantly, the PTC also includes the proprietary trading desks at major banks and investment dealers.

The PTC has proven extraordinarily efficient at locating pricing inefficiencies and quickly trading them away. Trading strategies typically involve rapid trading in and out of positions, with profits highly dependent on cost-efficient execution, which in turn is supported by the cost and efficiency advantages of electronic dealing platforms and prime brokerage arrangements. Despite initial resistance by some banks to opening up interbank dealing platforms to non-bank participants, both EBS and Reuters report very strong PTC demand for prime brokerage dealing relationships. As a result,

5. The top five banks by global market share were Deutsche Bank, UBS, Citibank, RBS, and Barclays Capital.

6. White-labelling describes an *electronic* transmission mechanism between the liquidity-supplying bank and the end user. However, many banks engage in *de facto* white-labelling by manually quoting clients a price taken directly off an interbank system (such as EBS or Reuters Dealing), simply adding a spread and passing along the liquidity supplied by other banks.

non-bank trading accounts are increasingly gaining access to the interbank market through prime brokerage channels.

The PTC has proven extraordinarily efficient at locating pricing inefficiencies and quickly trading them away.

Sixth, the efficiency of electronic price delivery has also allowed a specialized subsector of the B2C market, the retail aggregator platform, to develop. These electronic dealing portals cater to the smallest accounts, including households as well as small corporations, asset managers, trading firms, and institutions. Technology has so lowered the price of dealing in foreign exchange that some firms have found a profitable niche providing electronic services for foreign exchange dealing in retail amounts, typically defined as less than US$1 million (some retail platforms will open trading accounts for amounts as small as US$250). Retail aggregators capture the efficiencies in electronic price delivery and pass along very competitive pricing to retail accounts. As the costs of foreign exchange trading have dropped, retail participation in the market has surged. Some surveys suggest that retail accounts globally traded as much as US$60 billion a day in 2006; this number is projected to increase to well over US$100 billion a day by 2009 (Aite Group 2007).

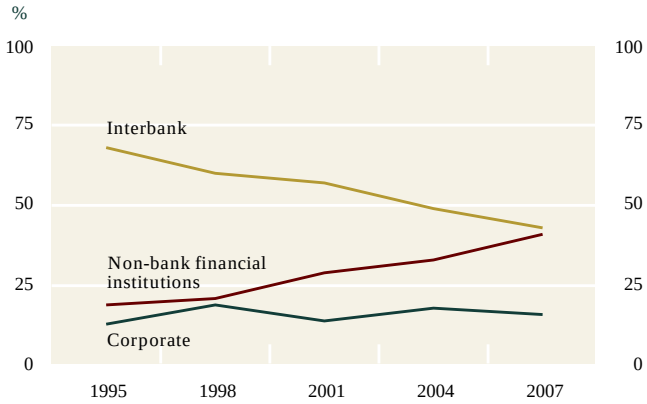
As a result of these various factors, the range of foreign exchange market participants has been broadening. This can be seen in Chart 3, which indicates that the proportion of overall trading volume accounted for by non-bank financial institutions, such as the PTC, institutional money managers, and retail aggregators, has been increasing, while the proportion accounted for by interbank trading has been declining. The proportion of activity accounted for by corporate accounts has been relatively stable.

Increasingly automated trading functions

Through the mid-1990s, most trading functions were done manually. This implied that dealers' market-making and proprietary trading activities were, in general, not systematically implemented. Traders had wide latitude when quoting prices to clients and other

Chart 3

Spot Foreign Exchange Flow by Organizational Type



Source: BIS and Reuters

banks, using their best judgment to manage the various risks generated by the market-making process and to assume proprietary risk positions. As a result, there was often little distinction between market making and proprietary trading, since market making involved warehousing imbalances (and thus price risks) in the client order flow, as well as trading in and out of the market in a continuous price-discovery process. Moreover, to manage the bank's position in the currency, traders were typically expected to have an opinion on the market and to express that bias when quoting prices to clients. Thus, a close relationship existed between agency trading functions (executing client orders) and principal functions (proprietary trading and the management of price risk). At many banks, one trader performed both functions.

With the technological innovations of the past decade, many trading functions once performed exclusively by traders are now increasingly performed by specialized computer programs. What distinguishes recent developments in this area is the use of application programming interface (API), the protocols that connect trading algorithms directly with the live price feeds on electronic trading platforms. With API, a trader can program the computer-based trading model to receive data from the market, process this information according to predetermined rules, and then generate buy and sell orders that are transmitted directly and immediately to the market without human intermediation.[7] The development of API has transformed all aspects of the trading process; specialized

7. Of course, human intermediation is required to reprogram the trading model for changing market circumstances or to override the algorithm's orders, if required.

computer programs now initiate trades, manage trade execution and order flow, and use complex algorithms to handle dealers' market-making functions.

Many trading functions once performed exclusively by traders are now increasingly performed by specialized computer programs.

API and algorithmic trading have had a marked impact on trading volume. The immediacy of computerized trading programs can produce a staggering number of trades, particularly for active-trading PTC accounts. Hedge funds already control a very large and rapidly growing pool of capital (estimated to be almost US$2.5 trillion in 2007 (Hedge Fund Intelligence 2007), and, like other PTC accounts, typically apply leverage to amplify their trading capacity. This extremely large pool of capital can then be rapidly traded through the market.

The combination of PTC penetration into the interbank market and computer-based trading has led to a surge in the proportion of algorithmically sourced foreign exchange volume. It is estimated that, since its introduction, algorithmic trading has achieved an approximate market share of 30 per cent on interbank platforms. Some analysts predict that algorithmic trading will eventually account for up to 70 per cent or more of foreign exchange volume, similar to what has occurred in equity markets (West 2007). The widespread availability of retail electronic trading portals and inexpensive computer power has enabled even smaller speculative accounts (such as day traders) to participate in the foreign exchange market.

A New, Hybrid Market

These three interrelated factors—electronic dealing platforms, a changing mix of market participants, and algorithmic trading—are rapidly changing the cost structure of the foreign exchange market. Reflecting this, foreign exchange is in a period of transition as global competition forces market participants to focus on areas where they have a clear comparative advantage. This is leading to a more distinct separation of principal and agency trading functions in the new paradigm foreign exchange market.

On one hand, a growing proportion of global market liquidity is supplied by large global dealers acting as *principals*, accepting and managing market risk for profit. Their client lists generate a significant order flow (especially through their PTC accounts), but these dealers will nonetheless warehouse temporary imbalances in the market and use client order flow to manage proprietary trading positions. In many respects, these top-tier dealers duplicate the market-making functions of local banks under the previous market structure, with one significant difference: their market-making activity is increasingly algorithmic in order to cope with the high speed and volume of modern foreign exchange markets. These market-making algorithms will often analyze client order flow and use the information gathered to guide the dealer's risk positioning, defining a form of automated, flow-based proprietary trading.

Foreign exchange is in a period of transition as global competition forces market participants to focus on areas where they have a clear comparative advantage.

On the other hand, the new paradigm market structure also contains significant *agency* elements, since many dealers also execute the orders of other market participants, but without assuming market risk themselves. There are two sources of agency operations that between them appear to account for a growing proportion of foreign exchange trading volume. First, an increasing proportion of the overall deal flow passing through the larger banks represents non-bank PTC accounts using prime brokered market access. Second, smaller (often domestic) banks that lack a comparative advantage in liquidity provision are increasingly outsourcing this function to larger banks via white-labelling or prime brokerage arrangements. Both prime brokerage and white-labelling are flow-based, low-margin commodity-like business models that succeed by keeping operating costs down, transactions volumes high, and exposure to market risk low.

As these agency functions grow in relative importance, the new paradigm foreign exchange market is adopting some characteristics of an "exchange" model. This

model already exists for several other asset classes—notably equities and commodities—where standardized financial products are traded on formal public exchanges by exchange members (e.g., the New York Stock Exchange). An exchange-based market model is no longer defined by the existence of a trading floor (many stock and commodity exchanges are moving towards electronic trading platforms). Rather, an exchange model has certain defining characteristics, including:

1. End-user accounts trade with each other through dealers (exchange members) who act only as their agents. Dealers do not act as principals by accepting or warehousing price risk, but provide only market access, credit-risk management, and other fee-based ancillary trading services.
2. This client-to-client (C2C) market is totally anonymous because end-user accounts trade though agents; there is no need for end-user accounts to know the identity of their ultimate counterparty as long as their agents (credit-risk managers) provide surety of settlement.[8]
3. With total trading anonymity and surety of settlement, all end users face the same price on the exchange without discrimination.

As the foreign exchange market evolves, some of the exchange-like characteristics are being integrated into its structure in several ways.

First, the relative importance of agency activities at many dealers is increasing as their business model expands to include fee-based market access, order execution and settlement, and credit-risk management for active trading accounts and other clients.

Second, although large dealers provide market-making functions, these market-making principal activities are increasingly provided by PTC accounts, both bank and non-bank. Dealers' proprietary trading desks have been likened to in-house hedge funds and clients of the dealers' agency-based order execution services. Prime brokerage allows non-bank PTC accounts to place orders in the interbank market like any other participant. In some aspects, the trading activities of the PTC—both bank and non-bank—replicate the functions of locals on the floors of commodity exchanges, or of specialists on some equity exchanges: absorbing temporary imbalances in the order flow and speculating on future price movements. In this manner, the ultimate source of liquidity in the market is more often a PTC account than a bank.

Third, prime brokerage arrangements provide anonymous trading for end-user PTC accounts, guarantee foreign exchange delivery to their clients, and give PTC accounts access to the same pricing as large global dealers. Surety of settlement is further reinforced because prime broker banks generally clear transactions among themselves through the CLS Bank.[9]

Fourth, not only have PTC accounts been increasing their access to the main interbank trading platforms, more recently, several electronic portals that cater to prime brokered PTC order flow have sprung up. Known as electronic communications networks (ECNs), some of these larger portals have been attracting trading volumes that are beginning to rival those of the main interbank platforms and multi-bank portals. (An example of an ECN is Hotspot FXi; its average daily trading volume is shown in Chart 4, below.[10])

Moreover, while trading volumes on these exchange-like ECNs have been building, an explicit exchange model has already developed in foreign exchange: the currency futures market at the Chicago Mercantile Exchange (CME). Recent (December 2006) estimates put average daily turnover volume for currency futures on the CME at US$80 billion. This rivals daily turnover on the main interbank portals and exceeds that on the major ECNs and multi-bank dealing portals. Moreover, volume growth on the CME is surging, since PTC accounts find the central clearing house exchange model well suited to their preferred trading strategies.[11] The CME's electronic trading platform also provides the high-speed API access and deep, liquid markets

8. Exchange members manage the credit risk of their clients.

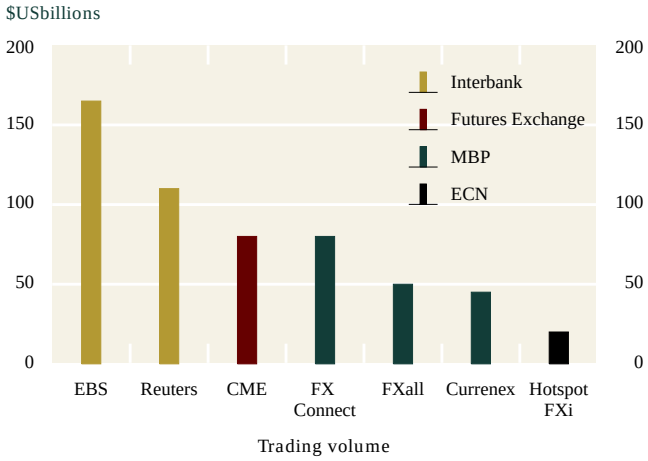
9. The CLS Bank is a clearing organization sponsored by the worlds' major central banks through which investment dealers can settle their trades. Surety of settlement is ensured through delivery-versus-payment protocols: banks must deliver their side of the trade to the CLS Bank before they will receive their counterparty's funds. More recently, and for a variety of cost and technical reasons, some banks have been exploring the possibility of netting trades between themselves outside of the CLS Bank.

10. Chart 4 also shows the average daily trading volume for Currenex, an electronic platform that combines modules from multi-bank portals and ECNs. FXall has recently introduced an ECN called Accelor, which operates separately from its multi-bank portal.

11. At the CME and other futures exchanges, all trades are settled through a central clearing house that provides total anonymity, surety of settlement, and non-discriminatory pricing for all counterparties.

Chart 4

Daily Trading Volume (end 2006)



Source: *FX Week*, 14 May 2007

that algorithmic trading routines depend on. As a result, the average daily trading volumes on the CME may be growing more rapidly and gaining on the traditional interbank trading platform.

While trading in currency futures contracts is not *spot* foreign exchange, it provides an efficient means of managing currency risk. Although this suits many PTC accounts, spot foreign exchange settlement is important to a broader array of market participants. Accordingly, to expand the exchange model, the CME recently entered into a joint venture with Reuters, named FXMarketSpace. This electronic trading platform duplicates the exchange-market features of a futures exchange, but in the spot foreign exchange market. Moreover, FXMarketSpace is open to all end-user accounts: the PTC, banks, institutional money managers, and corporations. This essentially creates a universal C2C trading space where dealers (exchange members) provide only market access and other agency-based ancillary trading services.

FXMarketSpace began trading early in 2007, and it is still too early to assess its overall impact on the spot foreign exchange market. But regardless of whether FXMarketSpace or ECN-based exchanges prevail—or even new trading platforms and protocols yet to be developed—the traditional paradigm of a geographically confined, relationship-based, bank-intermediated over-the-counter (OTC) market is being increasingly superseded by new participants, new business models, and new trading relationships that embody important elements of a global, C2C, exchange-style market.

Increased Liquidity in the New Market

Foreign exchange trading volumes have soared as the barriers to market access and the price of dealing in foreign exchange have declined. Not only are existing market participants trading more, but new participants are rapidly entering the market. The "democratization" of the foreign exchange market has resulted in the involvement of a much broader array of trading accounts. Indeed, much of the recent growth has come from PTC algorithmic traders, retail aggregators, corporate accounts, and institutional money managers who are increasingly treating foreign exchange as a separate, tradable asset class.

The larger trading volumes, increasing ticket numbers, and broader range of market participants have improved the liquidity of most currency pairs.

In addition to boosting trading volumes, this broader range of market participants has increased the diversity of opinion expressed in the market. Many of these new accounts trade in huge volumes but will split orders into a myriad of smaller deals spread throughout the trading day (growth in the number of trade tickets has exceeded growth in trading volume). By most measures, the larger trading volumes, increasing ticket numbers, and broader range of market participants have improved the liquidity of most currency pairs. For example, bid / offer spreads have been dramatically compressed; market flow is more evenly distributed across the trading day; order books are deeper (there are more resting orders at every price point); markets are more resilient to shocks; and both historical and implied volatility have recently been trending towards low levels.

Despite these measurable improvements, some have expressed concern over the growing PTC role in providing foreign exchange market liquidity because these accounts typically use highly leveraged, aggressive trading strategies. This may lead some speculative accounts to overcrowd similar positions, to overextrapolate existing price trends, or to inappropriately over-leverage price and credit risk. Such behaviour

has the potential to make the foreign exchange market less resilient or liquid under stressed conditions.

Although the foreign exchange market has not recently been tested by a period of severe stress, the increased price volatility observed in many global financial markets through the summer and autumn of 2007 was not accompanied by a marked deterioration in foreign exchange market liquidity. To the contrary, anecdotal evidence suggests that market participation by algorithmic PTC accounts, as well as overall foreign exchange trading volumes, increased remarkably during this period. By many accounts, liquidity remained deep in most foreign exchange markets even during the most volatile trading days of this extraordinary period.[12]

Whether broader PTC participation and the use of high-frequency trading algorithms have helped to moderate extreme price movements, or whether they may occasionally lead to price distortions and illiquidity in times of extreme market stress, remains an open question in foreign exchange and in many other markets. On balance, however, recent trading activity suggests that broadening the foreign exchange market has led to deeper liquidity, tighter pricing, more advanced trading technology, and more flexible, accommodative credit arrangements underlying market access (e.g., through prime brokerage). Moreover, foreign exchange markets have arguably become more operationally efficient as the automation of trading has dramatically lowered transactions costs. Allocative efficiency has also improved as technological innovation and new institutional trading arrangements have allowed trading risks to be unbundled, priced separately, and transferred to those more willing to bear them. This allows each market participant to manage those risks in which it has a comparative advantage.

As a result, the ongoing evolution in foreign exchange markets has benefited almost all market participants, not just the large investment dealers or the PTC.

Conclusion

The foreign exchange market is in a period of transition. Electronic trading platforms, algorithmic trading strategies, and a changing mix of market participants are driving market growth and accounting for an increasingly large share of global trading volume. In the process, the market structure that characterized foreign exchange through the mid-1990s is increasingly being displaced by a new structure that is different in several fundamental ways.

The lines of demarcation between buy-side and sell-side accounts; price-takers and price-makers; wholesale and retail trading platforms; and market players, banks, and the PTC are becoming increasingly blurred. So, too, are the distinction between exchange-traded and OTC transactions and those between bank-intermediated markets and disintermediated capital markets. In place of a primarily domestic market largely dominated by local banks, foreign exchange is moving towards an electronically linked international marketplace dominated by large global banks and non-bank trading funds, where all participants can access broadly similar pricing on a range of competing electronic platforms, and where a growing professional trading community increasingly supplies more of the liquidity and manages more of the price risks. Many banks are increasingly being disintermediated in the foreign exchange market and having to adjust their business models accordingly, making difficult choices as to where their core competencies lie: in an agency role providing fee-based ancillary trading services, or as risk-managing principals.

The distinctions between various financial markets have also been blurring as this new market model becomes increasingly multi-asset class in nature. Many PTC accounts have been moving into new asset classes, looking for new sources of return. The PTC's trading strategies have been generalizing to trade across different asset classes simultaneously, exploiting cross-price movements on multiple financial products. Real money institutional investors also have reasons to trade multiple asset classes simultaneously (for example, to buy foreign assets and hedge them at the same time, or to actively manage currency-overlay programs for their multi-asset-class portfolios).

The organizational structure of many banks is also evolving to reflect this new model. Departmental divisions are breaking down individual product silos in order to allow clients a more fully integrated multi-asset-class approach to transacting business. The skill set demanded of traders and dealers is changing as well: individuals well versed in high-value-added, multi-asset-class trading solutions are displacing single-product specialists who essentially provide foreign exchange price quotes on demand.

Similarly, electronic trading platforms are evolving to reflect the multi-asset-class approach increasingly

12. There were some reports of illiquidity in forward currency markets during this period, reflecting credit and technical concerns, but liquidity in spot foreign exchange markets remained firm.

demanded by influential PTC accounts. Several electronic trading portals—including many that originally focused on foreign exchange—are striving to integrate other financial products into their platforms, both by merger and acquisition, as well as by internal product development. Likewise, many public stock and commodity exchanges are not only moving away from trading floors to electronic platforms, but are also looking to bring a broader array of financial products into their organizations to help amortize the high fixed costs of advanced trading technologies.

As a result of these various inter-related, mutually reinforcing changes, foreign exchange (and other) markets are arguably becoming more open, transparent, and liquid. Operational efficiencies have also improved as trading costs have declined and innovations in risk management and broader market participation have allowed trading risks to be unbundled, priced more effectively, and dispersed more broadly.

Literature Cited

Aite Group. 2007. "Retail FX: Taking Center in Overall Market Growth." 23 July. Available at <http://www.aitegroup.com/reports/200707161.php>.

Bank for International Settlements (BIS). 2007. "Triennial Central Bank Survey of Foreign Exchange and Derivatives Market Activity in April 2007: Preliminary Global Results." Monetary and Economic Department. Available at <http://www.bis.org/publ/rpfx07.pdf>.

Euromoney Magazine. 2007. "Top Five Consolidate Lead in FX Market." July, pp. 165–78.

FX Week. 2007. "Market Nonchalant about FXMS Data." 14 May, p. 2.

Greenwich Associates. 2007. "Electronic Trading Systems Capture One-Half of Global FX Volume." 11 April. Available at <http://www.thetradenews.com/trading/foreign-exchange/772>.

HedgeFund Intelligence. 2007. "Global Hedge Fund Assets Surge 19% to $2.48 Trillion." Press Release, 1 October. Available at <http://www.hedgefundintelligence.com/images/590/55595/Global%20hedge%20fund%20assets%20$2.48trillion.pdf>.

NYSEData.com. 2007. "NYSE Group Volume in All Stocks Traded, 2007." Available at <http://www.nysedata.com/nysedata/Home/FactsFigures/tabid/115/Default.aspx>.

Price, T. 2007. "Consolidation: Threat or Opportunity for FX Execution Venues?" *e-FOREX Magazine*, January, p. 27.

Profit & Loss Magazine. 2007. "FX Just Keeps on Growing, Says Tower Group." July/August, p. 8.

Securities Industry and Financial Markets Association (SIFMA). 2007. "Average Daily Trading Volume in U.S. Bond Markets." Available at <http://www.sifma.org/research/pdf/Overall_Trading_Volume.pdf>.

West, K. 2007. "Algorithmic Trading: Behind the Trade." *International Banking Systems (IBS) Journal* (June) "Trading Platforms Supplement": 10–12.

The Effect of China on Global Prices

Michael Francis, International Department

- *Since China joined the World Trade Organization in December 2001, Chinese exports of consumer goods and imports of primary commodities have grown dramatically and are having major effects on the respective supplies and demands for these commodities.*
- *Globally, prices of consumer goods such as clothing, toys, and electronics are falling relative to other consumer goods and services, while the relative prices of commodities such as oil and metals have risen.*
- *Such relative price movements can have temporary effects on inflation, but monetary policy can adjust to keep inflation close to target over the medium term.*

Anyone walking through a local discount retailer during the past few years will have noticed an abundance of inexpensively priced goods, many of them labelled "Made in China." At the same time, gasoline prices have risen to levels not witnessed for more than a quarter of a century, apparently driven up by China's ever-rising demand for oil. From our drive to work to shopping on weekends, it seems that China is increasingly affecting our daily lives through the prices of the goods we buy. And Canada is not alone. Similar trends have occurred in all of the industrialized economies, suggesting that the "China price effect" is a global phenomenon.

This article examines the nature of these relative price changes and China's role in causing them. The analysis presented here suggests that, over the past five years or so, the growing supply of China's exports has exerted downward pressure on the prices of consumer goods, while its rising import demand has put upward pressure on global commodity prices. Yet although China appears to be affecting the prices of some goods relative to others, it is unlikely that the China price effect has had, or will have, a persistent effect on aggregate measures of inflation, such as core inflation, because inflation-targeting central banks have the tools to adjust policy to keep inflation close to target, thereby offsetting any persistent upward or downward inflationary pressure, regardless of its source.

China's Integration and Global Prices

Over the past 25 years, China's economy has grown at an annual average rate of 9.7 per cent. On a purchasing-power-parity (PPP) basis, the Chinese economy is now estimated to account for over 15 per cent of world gross domestic product (GDP), and it is predicted that per capita GDP in China will exceed US$8,000 by the

end of 2007.[1] By that measure, China is now richer than a number of East Asian rivals such as Indonesia and the Philippines. Much of this growth has been associated with an even more rapid increase in China's trade. Between 1980 and 2001, in current U.S.-dollar terms, exports grew at an annual average rate of just over 14 per cent, and since 2001, the annual growth rate of exports has approached 30 per cent. As a result, China is now the world's third-largest trading nation, with China's trade accounting for over 7 per cent of global trade.[2]

China's growth and integration into the global economy generally reflect the impact of its policies to foster growth and economic development.

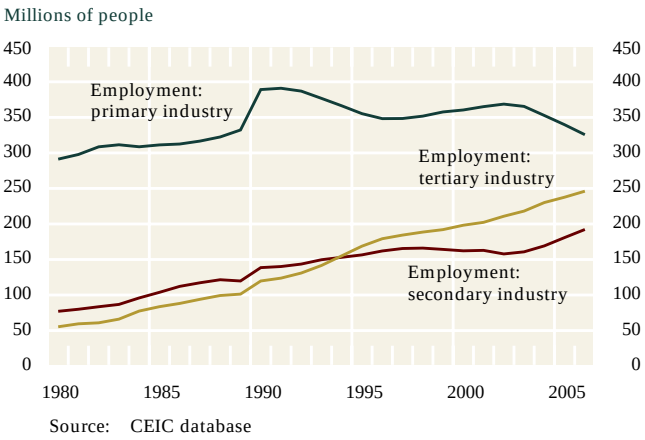
China's growth and integration into the global economy generally reflect the impact of its policies to foster growth and economic development. Before beginning its reforms in 1979, China had largely shut itself off from the rest of the world. In an effort to feed its large population, China's labour force was directed to work primarily in rural agriculture, while its exports were constrained by state plans and typically consisted of some key commodities, including oil, which were sold to earn the hard currency required to finance imports of capital equipment. During this period, China was extremely poor but largely self-sufficient. When reforms began in 1979, China set out gradually to transform the way it ran its economy. Instead of remaining inwardly focused, it elected to use trade as a means of drawing its abundant and underemployed labour out of agriculture to transform imported raw materials and inputs into manufactured goods for export as well as into new factories and infrastructure that have laid the foundation for growth. This process helped China to shift 200 million workers out of primary industries, such as agriculture, into secondary industries, such as construction and manufacturing, or the tertiary sector, which is mainly services (Chart 1).

In the first few years of the reform process, China was not large enough for its industrialization policy to have a significant impact on global markets. But in the past 10 years, and particularly the past 5, there is growing evidence that China has had important effects on the world economy through its contribution to world trade. In particular, as its trade has grown, China has become the dominant source of global supply and demand in some key sectors of world trade. Chart 2 shows that China's net exports now represent 39 per cent of world trade in semi-durable consumer goods, and 24 per cent of world trade in durable consumer goods, up from 26 per cent and 12 per cent, respectively, in 1995.[3] Likewise, China's imports account for 15 per cent of world trade in primary industrial supplies (unprocessed, non-food commodities) and 4.5 per cent of world imports of unrefined fuels and lubricants.[4]

Interestingly, as China's exports of consumer goods and imports of raw materials increased, the prices of these goods changed over time in a manner con-

Chart 1

China: Employment by Industry



Source: CEIC database

1. Because many (non-traded) goods and services are much less expensive in China than in the United States, simply converting China's GDP into U.S. dollars at the market exchange rate of 7.73 yuan per U.S. dollar (as at 1 April 2007) would understate the true size of the Chinese economy. A PPP conversion rate (2.095 for 2007) is therefore used for making international comparisons. Data on GDP, GDP per capita, and PPP conversion rates are from the International Monetary Fund *World Economic Outlook* database, April 2007.

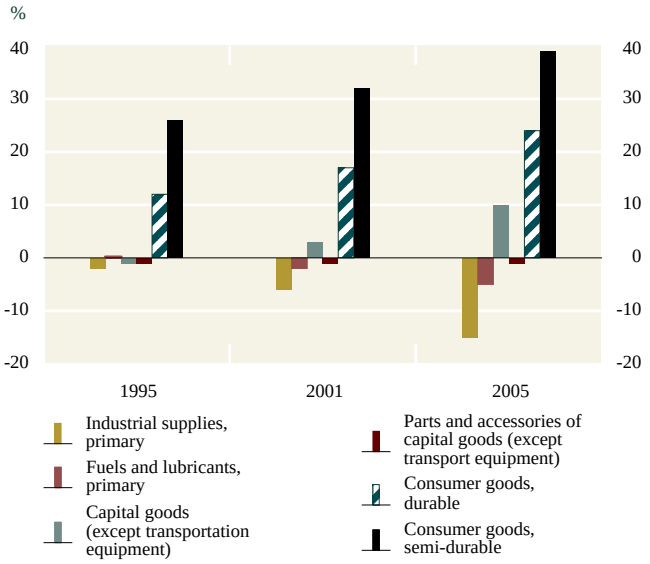
2. WTO Statistical database, 2006. Note that trade figures are calculated in U.S.-dollar terms, with conversion done at market exchange rates (not PPP-adjusted rates).

3. As a rule of thumb, durable consumer goods are manufactured household items that can be expected to last in excess of five years; and semi-durable goods those that last between two and five years. Televisions and household furniture are common examples of durable consumer goods, while clothing and children's toys are considered semi-durable consumer goods.

4. Figures are for China's net exports as a share of world trade in the specified industry.

Chart 2

China's Net Export Share in Global Trade (Selected Industries)



Source: United Nations Comtrade database

sistent with China's growing influence (Chart 3 and Chart 4). In particular, the prices of durable and semi-durable (DSD) consumer goods have fallen relative to the core consumer price indexes (CPI) in the industrialized countries, while world oil and metals prices have grown faster than U.S. CPI.

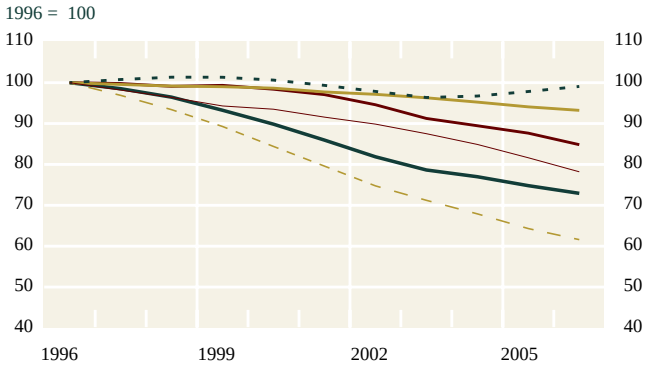
China has become the dominant source of global supply and demand in some key sectors of world trade.

However, this trend in relative consumer goods prices was present in the data even before China started to open its economy to world markets in 1979, and largely reflects the declining prices of goods relative to services in the CPI bundle. Baumol (1967) explained this trend by arguing that productivity growth was greater in the goods sector compared with services. As documented by Gagnon et al. (2003–2004), this was indeed the case in the major industrialized countries during the 1990s. But Gagnon et al. also point out that increased openness to international trade (especially from the emerging economies of Asia) is likely to have been playing an

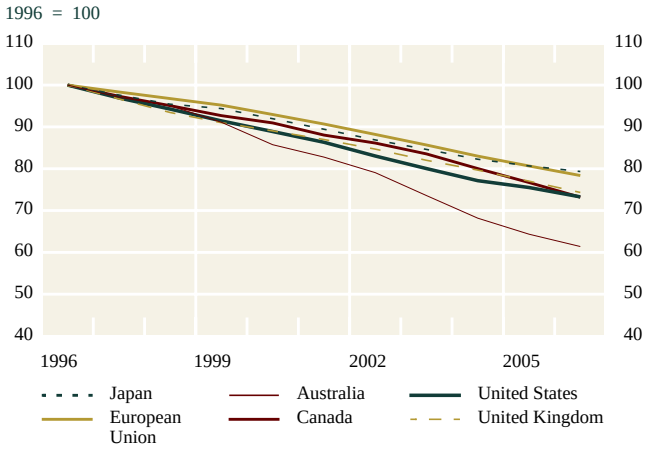
Chart 3

Price of Consumer Goods (Excluding Automobiles) Relative to Core CPI

Semi-durable goods



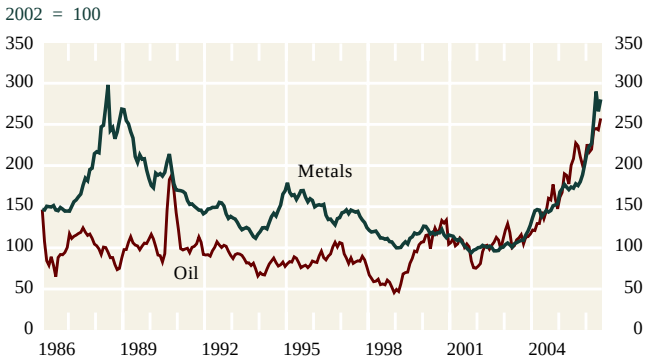
Durable goods



Source: National statistical agencies, Bank of Canada calculations

Chart 4

World Commodity Prices, Deflated by United States CPI



Source: International Monetary Fund *World Economic Outlook*, September 2006

important role in more recent years.[5] Given the rapid growth in the importance of China's trade over the past 5 to 10 years (as highlighted in Chart 2), the next two sections focus on this relatively recent development and its impact on the relative prices of consumer DSD goods and commodities.

China's Export Supply of Durable and Semi-Durable Consumer Goods

There are two main causes of China's dramatic increase in consumer goods exports — China's growth and its policy of trade liberalization. China's growth has been driven by several factors (Francis, Morin, and Painchaud 2005). Most studies have found that just under half of post-reform growth in China has resulted from improvements in total factor productivity (TFP), which reflects the efficiency with which China's resources are used. One important source of TFP growth is the reallocation of farm workers into more productive sectors such as manufacturing and services (Heytens and Zebregs 2003). Interestingly, as Chart 1 shows, the service-oriented tertiary sector has typically grown slightly faster than the industrial secondary sector. Beginning in 2003, however, employment in the secondary sector (including construction and manufacturing) has been growing rapidly (at 5 per cent, compared with 3.7 per cent before 2003) and appears to have drawn labour directly out of primary industries, mostly agriculture. Of the key secondary sectors, some of the greatest employment expansion has been in industries responsible for the production of DSD consumer goods. Increased employment in the electronics, garment, leather goods, and furniture industries, for example, accounts for more than one-third of the overall increase in employment in secondary industries (Table 1). Since the production of these types of goods is widely viewed as being labour intensive, this reallocation of labour could be expected to have resulted in a disproportionate increase in the supply of these goods on world markets.

Rapid capital accumulation, the other main source of China's growth, would normally tend to favour the production of more capital-intensive goods. In this instance, however, there are two important qualifications. First, in China, much of the investment is being allocated to construction and infrastructure development; and second, China's industrial policy (designed in part to create jobs for China's surplus agricultural labour) guided much of the remaining investment towards the labour-intensive export sector. The overall effect of growth on the output mix has therefore favoured the export-producing sectors over the import-competing sectors.

Table 1

Share of Total Increase in Employment in China's Secondary Industries, by Sector, 2000–06

	%
Electronics and communication equipment	17
Electric machinery and equipment	10
Garments and other fibre products	9
Textile industry	8
Leather and related products	7
Furniture manufacturing	3
Other	46

Note: Data are based on survey results, and coverage is incomplete. These statistics should therefore be considered indicative.
Source: CEIC database

Just under half of post-reform growth in China reflects the efficiency with which China's resources are used.

Trade liberalization also offers a good theoretical explanation for the increased supply of Chinese-produced DSD consumer goods. When China reduces import barriers, it frees up resources in protected sectors to flow into those exporting industries in which China has a comparative advantage. In addition, when industrialized countries reduce their import restrictions, it directly increases the supply of imported goods available in these markets from countries such as China. Thus, trade liberalization, on the part of both China and the industrialized nations would result in an increased supply of those goods in which China's comparative advantage is the greatest.

By the time China acceded to the World Trade Organization (WTO) in December 2001, many of the major obstacles to trade had been removed or eliminated as

5. Other explanations for the decline in consumer goods prices relative to the prices of services include difficulties in measuring quality improvements (especially for services); improvements in retail trade productivity—the so-called "Walmart effect" (Basker 2007); and an increasing demand for services compared with goods (owing to such factors as the aging of the population). On the commodity side, shocks to supply associated with weather and geopolitical tensions also play a role in affecting the relative prices of commodities.

Table 2

China: Applied Tariff Rates, 2001–05 (%)

	2001	2002	2003	2004	2005
Simple average applied rate	15.6	12.2	11.1	10.2	9.7
Industrial products	14.3	11.1	10.1	9.3	8.9
Motor vehicles	30.1	–	–	–	14.8
Textiles and clothing	21.1	17.5	15.1	12.9	11.5
Textiles	20.7	–	–	–	10.9
Clothing	24.1	–	–	–	15.8
Iron and steel	7.8	–	–	–	5.1
Agricultural products	23.2	17.9	16.3	15.0	14.6
Dairy	35.9	–	–	–	12.5
Grain	51.9	–	–	–	33.9
Oilseeds	32.0	–	–	–	11.1

Source: WTO (2006)

part of its export-led growth strategy. Nevertheless, upon accession, China undertook a number of important additional measures to further liberalize trade. These included significant tariff cuts (Table 2); the phasing-out of import quotas on such items as motor vehicles, petroleum products, rubber, iron, and steel; streamlining import-licensing requirements; and, importantly, removing restrictions on the right to trade. Before China's WTO accession, the right to trade was restricted to 35,000 Chinese enterprises. In some sectors, the right to trade was designated by the government, and import licences often prevented firms from reselling to the Chinese domestic market; instead, importing firms were required to use the imports for export.[6]

By 2005, 35 per cent of world trade in clothing and apparel was being sourced from China.

Although the key reductions in trade barriers resulting from joining the WTO occurred in China, accession also ensured that China's trade would be governed by the same set of agreements as other WTO members. This meant, for the most part, that Chinese-produced goods gained equal access to the markets of WTO members at the most-favoured-nation (MFN) tariff rates—that is, joining the WTO effectively levelled the playing field for Chinese goods that might otherwise have been discriminated against by the imposition of differential tariff rates.

Joining the WTO also meant that China benefited from the removal of quotas on textiles and clothing as negotiated under the WTO Agreement on Textiles and Clothing (ATC). China had been the world's largest exporter of clothing since 1995 (accounting for 22 per cent of world clothing trade at the time), but quotas in the major industrialized markets nevertheless severely limited Chinese clothing exports. As the quota system was gradually phased out after 2001, China's clothing exports accelerated, and by 2005, 35 per cent of world trade in clothing and apparel was being sourced from China.[7]

There are three main channels through which the quota reduction affected clothing prices in industrialized countries. First, the quotas caused the prices of Chinese clothing imports in industrialized countries to be higher than they would have been under free trade. As a result, over time, the import prices of these Chinese goods fell as the quotas were eliminated. Second, China is simply a cheaper source of clothing than many of its traditional competitors (such as Hong Kong or Turkey). Given the greater choice, consumers have been able to substitute towards the cheaper product. Third, increased competition from China has induced producers elsewhere to increase efficiency and reduce prices. Since clothing accounts for a substantial portion of the semi-durable consumption basket (almost 50 per cent in Canada and Europe, for example) these effects have likely had a significant impact on the prices of semi-durable goods in those industrialized countries that have phased out clothing quotas over the past five years.[8]

Effect on the relative prices of DSD consumer goods

Much of the problem with identifying a China effect on the prices of domestic consumer goods is the lack of readily available and reliable data on the prices and

6. Joining the WTO also had important legal and institutional consequences. Importantly, it bound the Chinese government to continue with the process of market liberalization. The effect of this commitment on businesses operating in China was probably important, but is difficult to quantify.

7. China has not fully benefited from the ATC quota phase-out because the United States and Europe maintained some quotas, which were extended past 2005 under safeguard clauses negotiated as part of China's entry into the WTO. These will expire at the end of 2007 in the case of Europe and 2008 for the United States.

8. The end of the system of quotas on clothing and textiles did not imply free trade in these goods. In the case of Canada, for example, the MFN tariff on clothing is typically 17 or 18 per cent.

quantities of Chinese-made goods in the advanced-country CPI baskets. Alternative approaches must therefore be used. One such approach is to try to "account" for the impact of China on import prices by breaking down the total import price (or, alternatively, the consumer price) into a share from China plus a share from the rest of the world (plus a share attributable to domestic production in the case of consumer prices). The effect of China on import prices then comes from three sources: the contribution of an increase in China's share (referred to as the import-penetration effect); the effect of changes in Chinese prices; and the indirect impact of increased competition from China on exporters in other countries and domestic producers (for consumer goods prices).[9]

Chinese goods are getting cheaper compared with similar DSD goods produced locally or abroad.

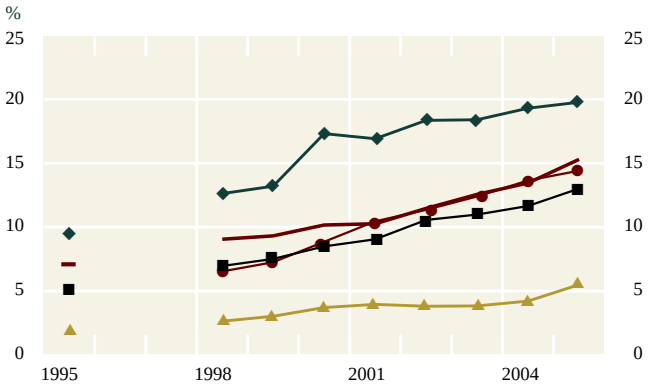
Using this accounting methodology, Nickell (2005) calculates that, between 2000 and 2004, China and India combined may have reduced CPI inflation pressure in the United Kingdom by just over half a percentage point per year through the import-penetration channel alone. For the United States, Kamin, Marazzi, and Schindler (2006) estimate that, between 1993 and 2002, the growing share of Chinese imports lowered import-price inflation by around 0.8 of a percentage point per year, translating into a small effect on consumer prices of around 0.10 of a percentage point per year. Pain, Koske, and Sollie (2006) find that, between 2001 and 2005, this import-penetration effect caused U.S. CPI inflation to be reduced by 0.12 of a percentage point in the United States and 0.13 of a percentage point in the euro area.

While the results are suggestive, there are limitations to the accounting approach. Importantly, for our purposes, the results do not identify the effect of China on DSD consumer goods relative to overall CPI. An indirect way of detecting a China effect on DSD prices is to examine the amount of expenditure on Chinese goods relative to total expenditure for a given category of consumer goods. If the imported good is a reasonable substitute for local goods, then the expenditure share of the imported good will rise as its price falls. Chart 5 shows that households in industrialized countries have been allocating an increasing share of their DSD consumption expenditures to Chinese goods, indicating that Chinese goods are getting cheaper compared with similar DSD goods produced locally or abroad, and over time are thereby contributing to downward pressure in core CPI goods.

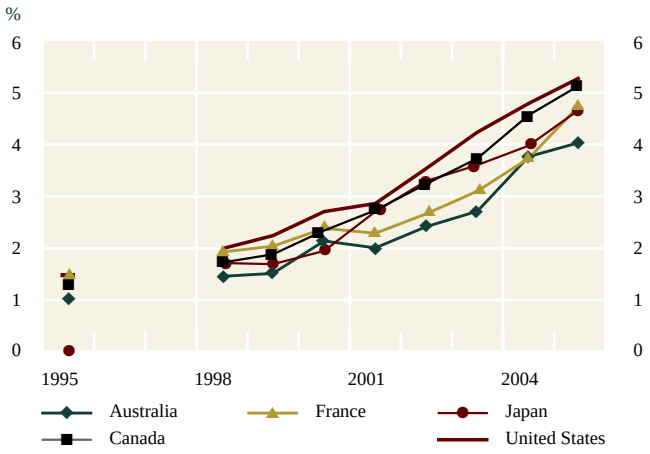
Chart 5

Share of Consumption Expenditure on Chinese-Made Goods, by Country

Semi-durable goods



Durable goods



Source: United Nations Comtrade database, national statistical agencies, and author's calculations

China's Import Demand for Oil and Metal Commodities

The flip side of China's export supply is its import demand. As discussed, China's industrialization policy

9. The direct effect of China on Canadian consumer prices is discussed in more detail in Morel (2007).

has created a demand for commodities and industrial supplies to be used in the production of its exports and for construction and infrastructure investment. Although China has a large resource sector, with over 10,000 mining enterprises employing five million people, in recent years, domestic output has been unable to keep up with domestic demand. This has created an import demand for commodities—especially primary commodities. This section looks at how China's demand for commodities is affecting the global prices of oil and metals.

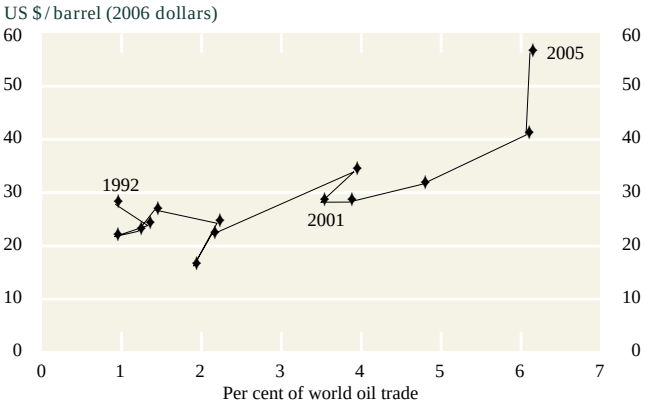
Effect on global oil prices

In the case of oil, between 2002 and 2004, China's oil consumption, driven by particularly rapid growth and a restructuring of its economy towards energy-intensive sectors, increased 28 per cent, or by approximately 1.5 million barrels per day (BP 2007). Consequently, China's share of world oil imports grew from approximately 3.5 per cent in 2001 to over 6 per cent of world oil trade in 2005.

During the 2002–04 period, although China's import demand was growing strongly, the impact on the world oil price was moderate (Chart 6). One reason was that the increase in China's demand at the time seems to have been perceived as temporary; hence, producers responded to what they thought was a short-term price rise by expanding production.[10] The U.S. Energy Information Administration (EIA), for example, predicted in its 2004 forecast that China's demand for oil for 2005 and 2010 would return to levels that it had predicted in its 2002 forecast.[11] Likewise, starting in 2002, the IMF consistently underestimated China's growth and did not significantly raise its projection of China's medium-term growth, from 8 per cent to 9 per cent, until 2006 (Chart 7). In response to these developments, global oil production rose, and spare capacity within the Organization of Oil Producing and Exporting Countries (OPEC) fell from an average of 3.7 million barrels per day between 1994 and 2002 to 1.5 million barrels per day between 2003 and 2005

Chart 6

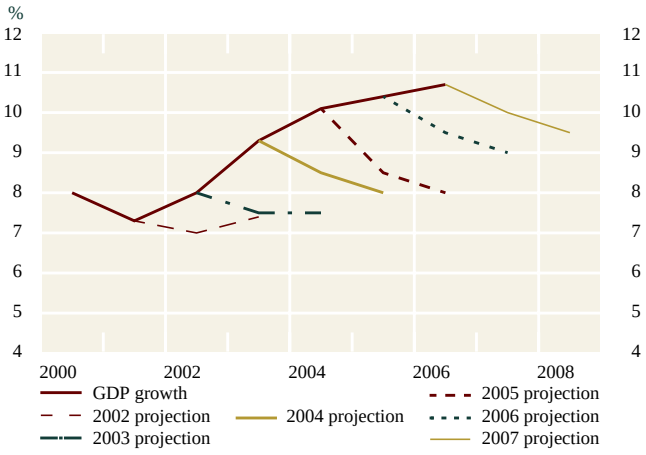
China's Crude Oil Imports (% of World Total) and the World Oil Price



Source: United Nations Comtrade database, IMF Commodity Price database, and author's calculations

Chart 7

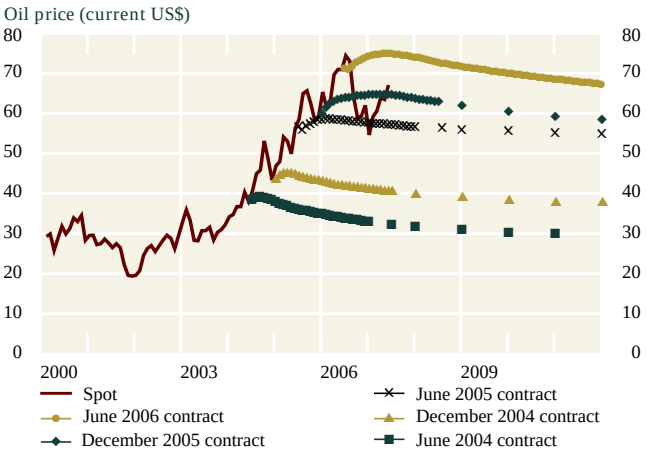
Growth of China's GDP and IMF Projections



Source: IMF *World Economic Outlook*, various years

Chart 8

Spot and Future Oil Prices, 2000–



Source: New York Mercantile Exchange (NYMEX)

10. Essentially, a resource producer can maximize the value of the resource by arbitraging across time: if the current price is higher than the expected future price, then the producer will increase output today to gain from the higher price. This response tends to dampen price movements caused by short-term changes in demand. Nevertheless, bottlenecks in the transportation, extraction, and refining stages resulting from the constraints on capital and labour mean that there is a limit to the arbitrage process, and commodity price rises may not be completely offset.

11. Rosen and Houser (2007) illustrate a similar point using the International Energy Agency's forecast for China's demand for oil in 2002.

(International Monetary Fund 2007).[12] It was these production responses that helped to moderate price rises at the time.

However, in 2006, at roughly the same time as the IMF began to make significant upward revisions to its outlook for China's growth, the EIA revised up its forecast for China's long-term oil consumption and then made another, more significant, upward revision in 2007, suggesting that the temporary increase in demand was now expected to be permanent. This change in expectations helps to explain why oil prices rose rapidly at this time.[13]

Around 2005–06, the evidence suggests that the market decided that what at first appeared to be a temporary increase in the growth of China's GDP and its demand for oil was likely to be permanent.

The impacts of the changing perception of China's demand for oil are nicely illustrated by oil futures market data (Chart 8). Prior to the middle of 2005, despite the growing demand from China, the oil futures market predicted that oil prices would fall from their spot prices (a phenomenon referred to as "backwardation") because of a belief in the market that the spot price at the time reflected a temporarily high level of demand relative to supply, which would be quickly alleviated, causing futures prices to fall. But, by the middle of 2005, the futures price curve had flattened out considerably, predicting that the oil price would remain close to US$55 per barrel. By early 2006, most of the backwardation in the futures price had been eliminated, with the oil price generally expected to rise from its spot price in the future.

Thus, starting around 2005–06, the evidence suggests that the market decided that increases in the growth of China's GDP and its demand for oil, which at first appeared to be temporary, were likely to be permanent. As a result, oil prices rose dramatically at a time when the increase in demand was deemed permanent and was exacerbated by already low capacity utilization.

To investigate the impact of this unexpected demand shock on global oil prices, Elekdag et al. (2007) use the Bank of Canada's Global Economy Model (BoC-GEM) to simulate the effect of an increase in East Asian productivity growth (in both traded and non-traded sectors) and energy intensity of oil usage.[14] They find that a surprise shock to China's oil demand could have pushed up oil prices by 20 per cent on impact and caused the long-run price to rise by as much as 60 per cent. Cheung and Morin (2007) use an econometric analysis to estimate the impact of emerging Asia on oil and metals prices. They find that there was a structural break in the data at the time of the 1997 Asian crisis—most notably for oil.[15] Since then, metals and oil prices, which had historically moved with the business cycle in the industrialized countries, have become increasingly aligned with emerging Asia's industrial activity. The results of these studies do not explain all the movement in oil and commodity prices, but they do suggest that China is having a significant and increasing effect on world commodity prices.

Effect on global metals prices

In many respects, China's impact on the world metals market is even more dramatic than its impact on the oil market. Between 2001 and 2006, metals prices almost tripled, with China accounting for more than 50 per cent of the increase in world demand for key metals such as aluminum, copper, nickel, and steel. Interestingly, though, with the exception of nickel, China managed to account for an even larger share of the increase in global refining capacity over the same period. In contrast, despite having a large mining sector, Chinese mine production failed to keep up (Table 3), implying that China was creating an excess

12. OPEC spare capacity refers to production capacity that can be brought online within 30 days and sustained for 90 days. Note that the period starting in 2003 was also one when the markets experienced numerous supply shocks, including the beginning of the war in Iraq—although the rundown in spare capacity significantly exceeded what was required to make up for forgone Iraqi production.

13. Theory suggests that the upward revision in future demand should produce an increase in long-run commodity prices, and that producers in the current period would have an incentive to conserve their supply to sell in the future. Thus, compared with a temporary demand shock, the supply response is slower and tends to produce an immediate and persistent increase in the commodity price. This increase in the long-run price, however, is also a signal to markets to increase investment in the extraction and refining sectors. Thus, the price rise may be partially offset over time by increased production capacity.

14. BoC-GEM is a sophisticated dynamic-stochastic general-equilibrium model of the world economy. For more details, see Lalonde and Muir (2007).

15. The authors attempt to identify whether the change in demand is coming from a trade channel or a growth channel. Unfortunately, given the limited amount of data in the post-2002 period, they have trouble deciphering between the two effects.

Table 3

China: Metals Demand and Supply, 2001–06

	China's share of world consumption and production growth (%)				China's domestic consumption and production growth (annual average) (%)			
	Alumi-num	Cop-per	Nickel	Steel	Alumi-num	Cop-per	Nickel	Steel
Refined consumption	59	54	58	60	20	10	23	19
Refined production	76	84	34	69	23	15	17	23
Mine production*	29	12	9	30	17	6	7	19

* Mine production for steel refers to iron ore production (in terms of metal content).

Source: World Bureau of Metal Statistics, except steel: International Iron and Steel Institute, Japan Iron and Steel Federation, United States Ecological Survey, and author's calculations

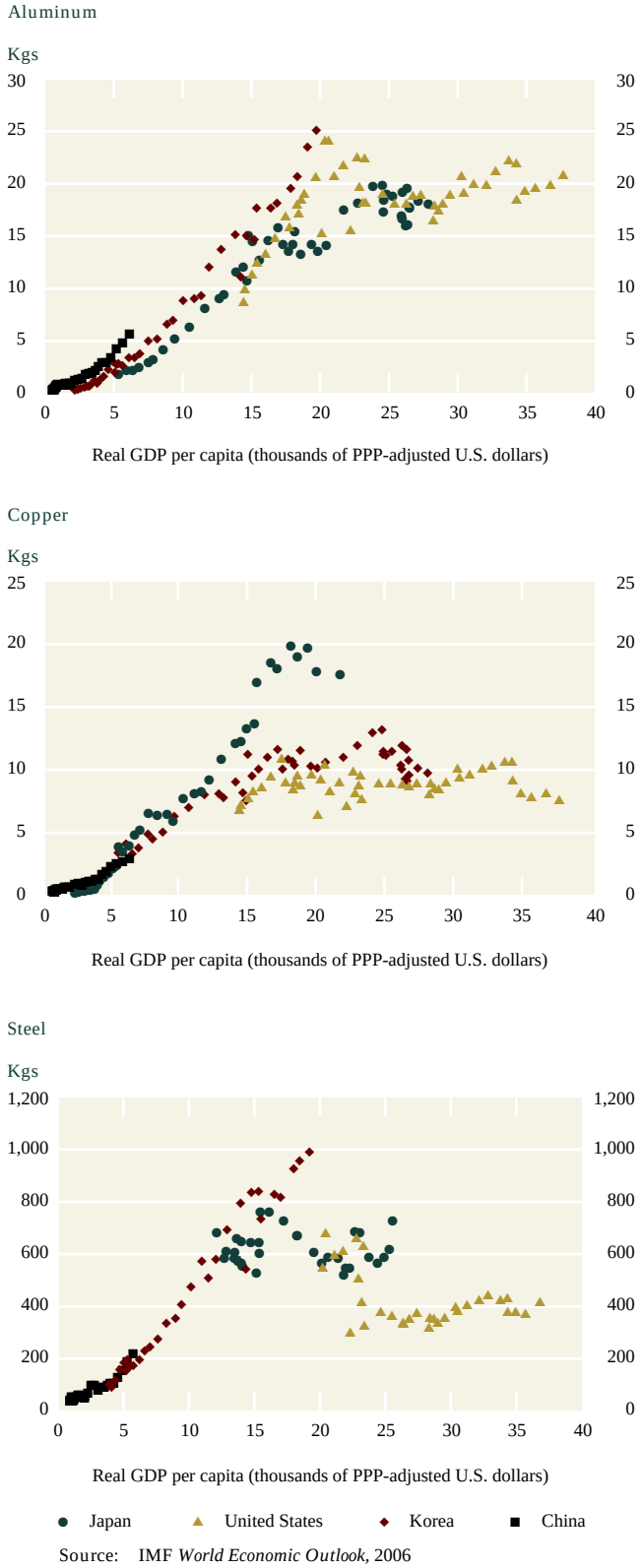
demand for unprocessed metals in the form of metal ores and concentrates.

On the consumption side, China's demand for metals is largely a function of its economic development. For an economy with per capita GDP approaching US$8,000, China's per capita demand for aluminum, copper, and steel is very much in line with that experienced by other countries at similar levels of development (Chart 9). Compared with advanced countries, however, China's demand for metals depends not only on the level of income, but also on how the metal is used. As Garnaut and Song (2006) argue, metals demand is particularly sensitive to the rate of urbanization and investment. For example, aluminum demand in China is more a function of domestic investment than household consumption when compared with industrialized countries. In 2002, construction accounted for 31 per cent of aluminum use in China, compared with 18 per cent in the advanced economies, whereas use in cans accounted for only 2 per cent of demand in China versus 12 per cent in advanced economies (Alcan 2004). Thus, in addition to its income growth, China's recent boom in construction and investment also helps to explain its commodity demand.

China's rapid increase in the production of refined metals is more difficult to explain (Rosen and Houser 2007). In part, it likely reflects laws and trade restrictions that limited the import of commodities, forcing domestic users to purchase locally produced metals, while firms producing for export had access to world markets. Although this practice has been gradually phased-out

Chart 9

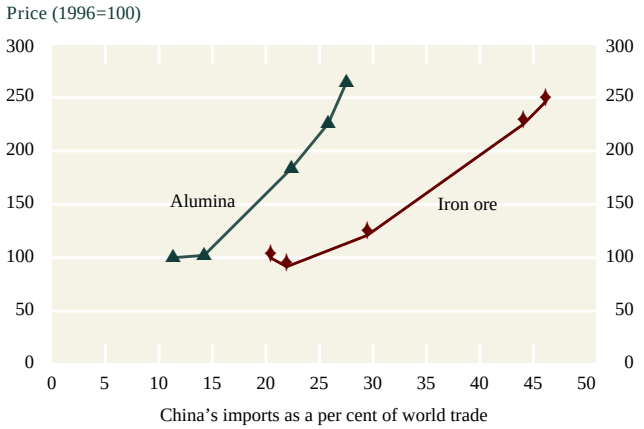
Per Capita Consumption of Metals



Source: IMF *World Economic Outlook*, 2006

Chart 10

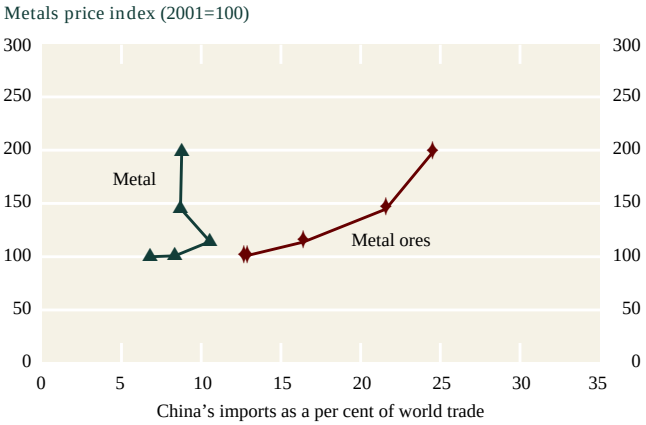
China's Impact on the Prices of Alumina and Iron Ore



Source: United Nations Comtrade database, IMF Primary Commodity Price database, and author's calculations

Chart 11

China's Import Demand and World Metals Prices, 2001–05



Source: United Nations Comtrade database, IMF Primary Commodity database, and author's caculations

since China acceded to the WTO, it remains applicable to the import of some commodities, such as iron ore (WTO 2006). In the case of aluminum, another factor that has encouraged the development of a refining industry is access to subsidized electricity (Alcan 2004). Low environmental standards and poor enforcement of those standards is also a problem in the Chinese metals-producing sector, although it can be a cost advantage (IOSC 2003, 2006). In the nickel-refining industry, for example, to save costs, Chinese firms have recently resorted to importing a low-grade ore (nickel laterite) used in the production of nickel pig iron—a substitute for nickel. Its use generates significant cost savings for Chinese firms, but the refining process produces high levels of pollutants (ABARE 2007; Lennon 2007). Whether China's advantage in producing refined metals persists as market forces penetrate the energy sector and demand for better environmental protection increases remains to be seen. In the interim, however, it seems likely that China will continue to be a major source of refining capacity.

Although growth in China's metals-refining sector has largely matched the rapid growth in consumption demand for refined metals, at just over 8 per cent, China's import share of the world refined-metals trade, while large, remained almost unchanged between 2002 and 2006. Yet growth in the demand for unrefined metal ores and concentrates has been dramatic: in 2002, China accounted for 13 per cent of world trade in metal ores; by 2005, it accounted for 25 per cent. Provisional estimates suggest that by 2006 it may have exceeded 30 per cent.[16] Not surprisingly, the prices of unprocessed metals such as alumina and iron ore have been highly correlated with China's import demand during recent years (Chart 10). Likewise, the increase in the world prices of refined metals have also shown a stronger association with the increase in China's import demand for metal ores and concentrates than with its demand for refined metals (Chart 11).

As with oil, although metal-producing resource companies likely anticipated China's rise and the associated increase in demand, the rate of increase in recent years and its permanence may have come as a surprise, creating an unexpected rise in world prices. In addition, given the scale of costs associated with developing new mining projects, it is likely that supply has been slow to adjust. As a result, mining companies have been straining to meet demand, and prices have risen and have remained elevated. In the longer term, since resource companies have revised their estimates of China's growth and the derived demand for metal ores upward, the capacity constraints ought to be grad-

16. China's growing and dominant share of world mineral ore markets is apparent in the market for alumina, where China's imports accounted for 12 per cent of world trade in 2002 and, by 2005, had reached 23 per cent; in the market for copper ore, where the corresponding figures are 14 and 23 per cent; and in the market for iron ore, in which China's import share of world trade had reached a staggering 46 per cent by 2005, up from 22 per cent in 2002. It is difficult to calculate figures for nickel, because of the significant impact of the substantial increase in the importance of cheap low-grade nickel ores during recent years, which tend to distort China's import figures. Ideally, a measure based on metal content could solve this problem, but such data are not readily available.

ually alleviated, and metals prices can be expected to fall. This future price decline is more likely for the most abundant metals, such as aluminum and iron, compared with such relatively scarce resources such as oil, where the scarcity value of the resource contributes a greater share to its price.

China's Effect on Global Inflation

In the long run, the rate of inflation in countries outside China is ultimately determined by monetary policy, not by China's impact on the relative price of oil or clothing. That said, central banks have had to contend with a series of large and persistent trade shocks emanating from China that may have surprised them as much as they have surprised the IMF or mining companies. Given the lags associated with monetary policy actions, these shocks can therefore be expected to have an effect, albeit temporary, on measured inflation.

Nevertheless, it is not well understood how China's trade, which is thought to act on real variables and relative prices, affects nominal prices and measured inflation. Possible channels include the weighting of goods in the CPI basket—the impact of some relative prices (such as DSD consumer goods) on measured inflation may be more important than others (such as energy); an effect on firms' price-setting behaviour owing to increased import penetration associated with competitively priced imports of Chinese DSD consumer goods; downward pressure on wages, which reduces production costs; and upward pressure on production costs stemming from higher commodities prices.

Given China's size, and the rapid acceleration in its trade following its WTO accession, it is therefore not surprising that some observers (such as Nickell 2005) see China as playing a significant role in the inflationary process, at least over the short term. At this point, however, given that the relative price effects on DSD goods and commodity prices are somewhat offsetting, and that monetary policy in most industrialized countries is forward looking and aimed at price stability, definitive empirical evidence that China is a net source of disinflation (or inflation) remains elusive.

Conclusion

This article has explored the role that China plays in determining global prices and, in particular, the effect that China is having on the relative prices of consumer goods and commodities via its export supply and import demand for these goods. The evidence suggests that, following its accession to the WTO, China played a significant role in restructuring global trade and hence affected relative prices. In the market for clothing, the phase-out of quotas seems to have significantly increased the availability of inexpensive clothing from China. As a result, the global clothing market has become more competitive, clothing prices have fallen, and expenditure on imported Chinese clothing has risen. A similar pattern seems to be present in other DSD consumer goods markets. In terms of commodity imports, China's economic and trade developments appear to have grown much faster than expected, causing a larger-than-anticipated increase in global demand for oil and metals. Together, these two effects help to explain the recent change in the relative prices of these goods.

Looking forward, China continues to have a large supply of labour in the primary sectors of its economy, which can be expected to continue its migration into the DSD consumer goods sector for some time. This process will help to keep downward pressure on the relative prices of these goods.

China's demand for oil and metal commodities can be expected to grow strongly for a number of years to come.

For oil and metals, China's size and growth are likely to remain among the key factors driving the growth of global demand for some time. By way of comparison, in per capita GDP terms, China's current level of development is not dissimilar to that of Japan in the early 1960s. Thus, if Japan's experience is relevant, China's demand for oil and metal commodities can be expected to grow strongly for a number of years to come. To keep markets in equilibrium, either prices will rise or the supply side will adjust. History suggests that supply does adjust, but that the adjustment will be slow, given the scale of the required adjustment, the lags involved in establishing the necessary capacity, and caution surrounding the risks associated with increases in demand being driven by a single market. Hence, the relative prices of commodities can also be expected to remain somewhat elevated.

Finally, although this article has focused on the effect of China, the emergence of other labour-abundant emerging economies, such as India, which are likely to follow in China's footsteps, must be acknowledged. There will no doubt be some differences in how these economies affect global markets and prices, but China's rise suggests that the relative price effects could be significant.

Literature Cited

ABARE. 2007. *See* Australian Bureau of Agricultural and Resource Economics.

Alcan Inc. 2004. *China's Aluminum Industry: Setting the Record Straight*.

Australian Bureau of Agricultural and Resource Economics (ABARE). 2007. "Australian Commodities, June Quarter."

Basker, E. 2007. "The Causes and Consequences of Wal-Mart's Growth." *Journal of Economic Perspectives* 21 (3): 177–98.

Baumol, W. J. 1967. "Macroeconomics of Unbalanced Growth: The Anatomy of an Urban Crisis." *American Economic Review* 57 (3): 415–26.

BP. 2007. *BP Statistical Review of World Energy*, June.

Cheung, C. and S. Morin. 2007. "The Impact of Emerging Asia on Commodity Prices." Bank of Canada Working Paper. Forthcoming.

Elekdag, S., R. Lalonde, D. Laxton, D. Muir, and P. Pesenti. 2007. "Oil Price Movements and the Global Economy: A Model-Based Assessment." Bank of Canada Working Paper No. 2007-34.

Francis, M., F. Painchaud, and S. Morin. 2005. "Understanding China's Long-Run Growth Process and its Implications for Canada." *Bank of Canada Review* (Spring): 5–17.

Gagnon, E., P. Sabourin, and S. Lavoie. 2003–2004. "The Comparative Growth of Goods and Services Prices." *Bank of Canada Review* (Winter): 3–10.

Garnaut, R. and L. Song, 2006. "China's Resources Demand at the Turning Point." In *The Turning Point in China's Economic Development*, 276–93. Canberra: Asia Pacific Press at the Australian National University Press.

Heytens, P. and H. Zebregs. 2003. "How Fast Can China Grow?" In *China: Competing in the Global Economy*, edited by W. Tseng and M. Rodlauer, 8–29. Washington, D.C.: International Monetary Fund.

Information Office of the State Council of the People's Republic of China (IOSC). 2003. "China's Policy on Mineral Resources." Government White Paper. Available at <http://www.china.org.cn/e-white/index.htm>.

———. 2006. "Environmental Protection in China (1996-2005)." Government White Paper. Available at <http://www.china.org.cn/e-white/index.htm>.

International Monetary Fund. 2006. "The Boom in Nonfuel Commodity Prices: Can It Last?" *World Economic Outlook* (September): 139–70.

———. 2007. "Global Prospects and Policy Issues." *World Economic Outlook* (April): 1–46.

Kamin, S. B., M. Marazzi, and J. W. Schindler. 2006. "The Impact of Chinese Exports on Global Import Prices." *Review of International Economics* 14 (2): 179–201.

Lalonde, R. and D. Muir. 2007. "The Bank of Canada's Version of the Global Economy Model (BoC-GEM)." Bank of Canada Technical Report No. 98.

Lennon, J. 2007. "The Chinese Nickel Outlook and the Role of Nickel Pig Iron." Presentation to International Nickel Study Group, 11 May. Available at <http://www.insg.org/presents/Mr_Lennon_May07.pdf>.

Morel, L. 2007. "The Direct Effect of China on Canadian Consumer Prices." Bank of Canada Discussion Paper No. 2007–10.

Literature Cited (cont'd)

Nickell, S. 2005. "Why Has Inflation Been so Low since 1999?" *Bank of England Quarterly Bulletin* 45 (1): 92–107.

Pain, N., I. Koske, and M. Sollie. 2006. "Globalisation and Inflation in the OECD Economies." Organisation for Economic Co-operation and Development (OECD) Economics Department Working Paper No. 524.

Rosen, D. H. and T. Houser. 2007. "China Energy: A Guide for the Perplexed." A Joint Project by the Center for Strategic and International Studies and the Peterson Institute for International Economics. May.

World Trade Organization (WTO). 2006. "Report by the Secretariat: People's Republic China." Trade Policy Review.

Estimating the Cost of Equity for Canadian and U.S. Firms

*Lorie Zorn, Financial Markets Department**

- *There has been a concern among policy-makers that the cost of equity financing may be higher in Canada than in the United States, but the empirical evidence supporting this view is mixed.*
- *We improve on previous studies by implementing a forward-looking, firm-specific approach to estimating the nominal cost of equity for Canada and the United States that controls for firm characteristics, industry effects, and business cycle effects.*
- *We find that greater firm size and greater liquidity of a firm's stock are associated with a lower cost of equity, while greater firm leverage and greater dispersion in analysts' earnings forecasts are associated with a higher cost of equity. Moreover, we find that higher yields on longer-term sovereign bonds increase a firm's cost of equity.*
- *After taking firm-level and aggregate-level factors into account, the cost of equity was approximately 30 to 50 basis points higher in Canada than in the United States over the 1988–2006 period as a whole, but this differential appears to be lower in the post-1997 period.*

Financing costs are important for both firms and the economy, affecting investment decisions and, ultimately, economic growth. Since equity is an important component of a firm's financing structure, Canadian firms may not undertake as many projects that could potentially enhance growth if the cost of equity financing in Canada is relatively high. Considering the overall size of the equity stock in Canada, even small differences in the cost of equity financing can have a substantial impact.[1]

The cost of equity, which can be defined as the return expected on a firm's common stock, represents the compensation demanded by shareholders for providing capital and assuming the risk of waiting for this return.[2] Thus, in addition to the risk-free return, the cost of equity incorporates an equity-risk premium—the incremental payoff from holding a risky equity security rather than a risk-free security.

There has been a concern among policy-makers that financing costs may be persistently higher in Canada than in the United States. The Capital Markets Leadership Task Force begins its 2006 report, for example, with the premise that the cost of capital in this country needs to be reduced for Canadian firms to compete effectively with those in the United States (Boritz 2006). Similarly, the report of the Task Force to Modernize Securities Legislation in Canada (2006) reinforces the notion of a "made-in-Canada" risk premium that

* The research reported in this article is summarized from a working paper written by Jonathan Witmer and the author (Witmer and Zorn 2007).

1. As of 31 December 2006, the market capitalization of the Toronto Stock Exchange (TSX) was just over $2 trillion. During 2006, TSX firms raised over $41 billion through share issues. Available on the TSX website at <http://www.tsx.com>.

2. The cost of equity can be expressed in real or nominal terms, depending on whether real or nominal returns per share are used in its estimation.

increases the cost of equity capital in Canada and discounts the trading price of Canadian shares.[3]

The empirical evidence supporting this view is mixed. Multi-country studies indicate that the costs of equity for Canada and for the United States are comparatively close on a worldwide scale. The magnitude and relative ranking of these estimates vary across studies, however. Claus and Thomas (2001), for example, calculate a cost of equity for Canada that is 20 basis points (bps) lower than that of the United States.[4] The frequently cited results of Hail and Leuz (2006) indicate a cost of equity for Canada that is 30 bps greater than that of the United States.[5]

Policy efforts aimed at fostering a healthy environment for investment financing in Canada can be enhanced by a better understanding of the drivers of the cost of equity.

Canadian policy-makers have an interest in fostering a healthy environment for investment financing and, in the end, economic growth in Canada. Policy efforts can be enhanced by a better understanding of the drivers of the cost of equity in Canada, particularly compared with those of other countries.

This article presents estimates of the influences on the cost of equity in Canada and the United States using an updated methodology that controls for firm characteristics and aggregate-level factors. We begin with a brief review of the empirical literature. Next, we summarize the key factors that affect the cost of equity. We then present a comparison of Canadian and U.S. firms. Finally, the contributions of key factors to the cost of equity for Canadian and U.S. firms are quantified and discussed, along with implications for policy-makers.

Estimating the Cost of Equity

Only a handful of studies over the past 15 years have estimated a cost of equity for Canada, and the results vary. The studies not only disagree on the size of Canada's cost of equity, with estimates ranging from 5.4 per cent to 10.8 per cent, but they also disagree on how Canada compares with the United States. Some estimate a slightly higher cost of equity in Canada; some estimate that Canada's cost of equity may be 3 per cent lower.[6]

Why has the empirical literature failed to provide solid conclusions? One likely reason is that only recently has a true forward-looking, firm-specific approach to estimating the cost of equity been applied to Canada. Because sufficient firm-level data were not available before the mid-1990s, most estimates are based on realized, market-level returns on stocks and sovereign bonds. Typically, the methodology used in these studies estimates a constant equity-risk premium based on the differences in nominal returns earned on equities and bonds during a lengthy period of time (often 50 years or more). Because of historically lower stock market returns and higher bond yields in Canada relative to the United States, these studies have tended to find a lower equity-risk premium for Canada. Although risk-free rates have tended to be slightly higher in Canada, the result is often a lower cost of equity for Canada relative to the United States.[7] However, the period over which this market-level risk premium is calculated can lead to very different cost-of-equity results.

In addition, research to date has not been focused on making a thorough comparison between Canada and the United States. Rather, the cost of equity has often been estimated as a preliminary step to answering other questions (such as whether differences in a country's legal environment have an impact on the cost of equity). These country cost-of-equity estimates typically do not account for firm-specific characteristics and aggregate-level factors that could affect the cost of equity. Differences across these studies could therefore be attributed to the different characteristics of individual firms in each sample. In addition, variations in the estimates might be exacerbated by using a relatively small sample of firms in Canada compared with the United States.

3. The report cites the findings of Hail and Leuz (2006) and King and Segal (2003, 2006).

4. They estimate that Canada's cost of equity is 10.8 per cent over the period 1985–98, compared with 11 per cent for the United States.

5. Hail and Leuz estimate a cost of equity for Canada of 10.5 per cent over the period 1992–2000, versus 10.2 per cent for the United States.

6. See Witmer and Zorn (2007) for a discussion of the empirical literature.

7. See, for example, Booth (2001); Jorion and Goetzmann (2000); and Hannah (2000).

Lastly, although the cost of equity is, by definition, linked to the risk-free rate, it may also be insightful to consider the interest rate environment and how this affects the financing costs of individual firms in Canada.

Using information from stock prices and stock analysts' forecasts of firm earnings, we estimate a nominal cost of equity for Canadian and U.S. firms, then compare these estimates.

We address all of these issues by employing a methodology that uses information from stock prices and stock analysts' forecasts of firm earnings to estimate a nominal cost of equity for each firm.[8] Our cost-of-equity estimates are intuitively appealing because they reflect expected future returns to shareholders: in this approach, the cost of equity is the rate of return that sets the current stock price equal to the present value of expected future cash flows to shareholders. We compare these estimates for Canadian and U.S. firms over the 1988–2006 period, first at a broad level, and then controlling for firm characteristics, industry effects, and business cycle effects in a panel regression analysis. As an additional step, we examine the impact of longer-term sovereign bond yields (a proxy for the risk-free rate) on these cost-of-equity estimates.

What Drives the Cost of Equity?

A firm's cost of equity can be affected by several factors, which can be classified both at a firm level and at a broader level. Generally, the more these variables increase the perceived riskiness or uncertainty of future returns to shareholders, the more shareholders will demand to be compensated for this risk, and the higher will be the firm's cost of equity. Because our analysis incorporates these variables, it is important to establish their expected effect on a firm's cost of equity in order to help interpret our results:

- *Firm size*: Since there is usually more information regarding the management and potential earnings of larger firms, the uncertainty regarding the future returns of such firms is reduced. Thus, we would expect a firm's cost of equity to be negatively related to its size.
- *Financial leverage*: Given that payments to debt holders have priority, an increase in debt (or greater financial leverage) and fixed interest costs will make returns to equity holders more sensitive to changes in earnings (i.e., more risky). Thus, we would expect greater financial leverage to increase a firm's cost of equity.
- *Corporate taxes*: Corporate taxes have an indirect effect on the cost of equity by reducing the impact of financial leverage. Since interest payments on debt are tax deductible, corporate taxes reduce the effective cost of debt. So where corporate taxes are levied, leverage provides a risk-less tax shield, such that the overall risk of the firm is lower for the same amount of financial leverage. Through this link with financial leverage, we would expect the cost of equity to be negatively related to corporate taxes.
- *Stock liquidity*: Investors require extra compensation to cover the costs of buying and selling a security. These transactions costs tend to be lower for more frequently traded or more liquid stocks.[9] Thus, we expect firms with greater stock liquidity to have a lower required return and, hence, a lower cost of equity.
- *Forecast dispersion*: Investor uncertainty regarding future returns could grow with the variability and reduced accuracy of analysts' earnings forecasts for a firm. Thus, we would expect greater disagreement or dispersion in analysts' forecasts to increase the cost of equity.

In addition to these firm-specific characteristics,[10] other factors affect the cost of equity at a broader level:

8. See Witmer and Zorn (2007) for details on our methodology, including potential shortcomings.

9. Securities regulation and competition between trading platforms or exchanges have an impact on average stock liquidity as well.

10. Although not included in our analysis, ownership structure may also affect a firm's cost of equity. King and Santor (2007) find that Canadian firms with dual-class shares have a lower equity valuation than those firms with non-dual-class shares. Given the inverse relationship between a firm's cost of equity and its share price, this implies a higher cost of equity for firms that use dual-class shares.

- *Industry factors*: Certain cost-of-equity drivers will be common across firms in the same industrial group. For example, industries such as mining will have a high proportion of fixed costs. This higher operating leverage will cause profits to be more sensitive to changes in revenue, thus increasing the riskiness of returns to the firms' shareholders and the cost of equity in these industries. We attempt to capture industry-wide effects on the cost of equity by including industry dummy variables in our analysis.
- *Economic conditions*: Studies have shown that expected returns for equity markets tend to be countercyclical; i.e., they are lower under strong economic conditions and higher under weak economic conditions. Thus, we expect business cycle effects on the cost of equity as well and include dummy variables for each year in our sample period to account for this.

Differences in the cost of equity across firms can also be affected by such variables as the degree of financial market segmentation, unexpected movements in exchange rates, inflation uncertainty, differences in personal taxes, and different legal and regulatory environments, including enforcement. Because our focus is on firm-level drivers of the cost of equity that can easily be represented, we do not address these other factors. (Although other studies have examined the relationship of some of these factors with the cost of equity, none has comprehensively included all of these variables.) An analysis of some of these other effects is planned in future work, however, and this might shed further light on the cost of equity for Canadian firms.

Empirical Results

Canada-U.S. comparison

Given the factors affecting the cost of equity, it is interesting to first compare Canadian and U.S. firm characteristics. Taking a sample of firms over the period 1988 to 2006,[11] tests are performed to determine whether there are differences between the Canadian and U.S. median for the five identified firm characteristics (Table 1). The tests indicate that, compared with U.S. firms, Canadian firms in our sample are smaller, have a lower effective tax rate, a higher amount of debt in their capital structure, a lower stock turnover (a proxy for stock liquidity), and a higher dispersion of forecasts among analysts. When the cost of equity is estimated for each firm and year, we find that the median cost of equity is 11.5 per cent for Canadian firms, compared with 10.9 per cent for U.S. firms over the 1988–2006 period.[12] Given the differences in firm characteristics, it is not surprising that the median cost-of-equity estimate for firms in the Canadian sample is higher than that for firms in the U.S. sample.[13] As such, it is important to control for these firm-level differences in order to make a relevant comparison across countries.

Table 1

Sample Statistics for Canadian and U.S. Firms, 1988–2006

	Canada	United States	Median difference
Size (total assets)	US$364.2 million	US$446.8 million	-US$82.7* million
Financial leverage	0.36	0.33	0.03*
Taxes	0.35	0.36	0.01*
Stock liquidity	0.30	0.94	-0.64*
Forecast dispersion	0.06	0.03	0.03*
Cost of equity	11.49	10.86	0.64*

* Significant at 1 per cent

Notes: Size is calculated using book values from Compustat and is converted into U. S. dollars. Financial leverage is calculated as the ratio of long-term debt to equity using book values. Tax is the ratio of income taxes to pre-tax income and is restricted to a range between 0 and 1. Stock liquidity is proxied by turnover and is the number of shares traded in the previous year divided by the total number of shares outstanding in Compustat. Forecast dispersion is the cross-sectional standard deviation of analysts' earnings forecasts denominated in U.S. dollars. The nominal cost of equity is based on forecasted earnings from I/B/E/S and is calculated using the average of four different forward-looking models.

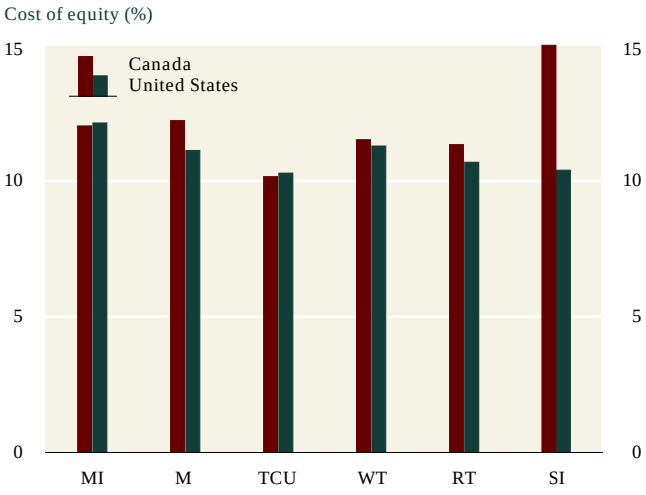
11. Our sample contains Canadian and U.S. non-financial firms covered by the Institutional Brokers Estimate System (I/B/E/S) and Compustat. After merging the two datasets, we have 3,419 Canadian and 31,005 U.S. observations.

12. We use an average of four forward-looking models to estimate the nominal cost of equity. For more details, including robustness to different assumptions, see Witmer and Zorn (2007).

13. Our cost-of-equity estimates are likely higher than those from previous studies because our sample includes more small firms.

Chart 1

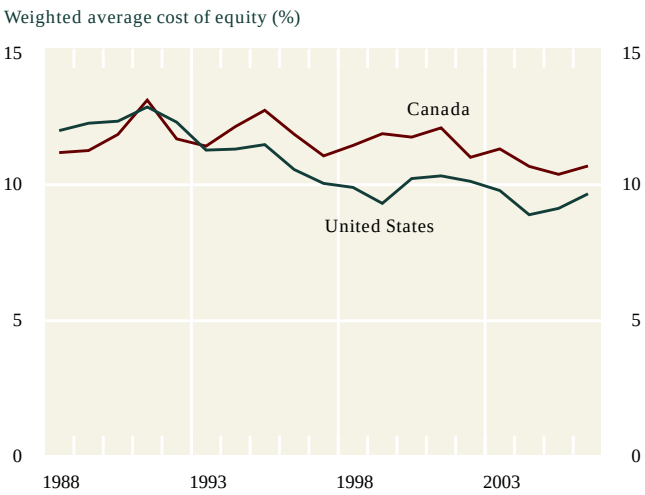
Median Cost of Equity by Industry, 1988–2006



Note: MI = Mineral industries, M = Manufacturing, TCU = Transportation, communication and utilities, WT = Wholesale trade, RT = Retail trade, and SI = Service industries.

Chart 2

Cost of Equity by Year, 1988–2006



Note: The cost of equity is weighted by firm size for Canada and the United States on a yearly basis. Early in the period, when the Canadian sample is dominated by large firms, Canada's cost of equity is smaller. Over time, the median firm size for Canada falls as the proportion of smaller firms rises. Other factors that affect differences in firms' cost of equity are not controlled for here.

It is not surprising that the median cost-of-equity estimate for firms in the Canadian sample is higher than that for firms in the U.S. sample.

Next, we break out industry and business cycle effects, presenting the cost-of-equity estimates by industry grouping (Chart 1) and by year (Chart 2). Grouping firms by their two-digit Standard Industry Classification code, it appears that Canada has a higher cost of equity in four of the six broad industry groups (although, again, we are not at this point controlling for all of the aforementioned firm characteristics). Looking at Chart 2, some general observations can be made: there is a downward trend in the cost of equity for both countries; there are similar cycles in the cost of equity for Canada and the United States; and Canada appears to have a higher cost of equity for most of the period. This reinforces the notion that the cost of equity is not static, but time varying. However, there are also differences in our sample of firms across time and countries. For example, at the beginning of the period, the Canadian sample is dominated by larger firms, but the median firm size falls over time as the proportion of smaller firms rises. In contrast, the median firm size in the U.S. sample increases significantly over time. Because of sample differences such as this, there is a need to incorporate all of the identified factors into our analysis before making conclusions about the relative cost of equity.

Regression analysis

A regression analysis (see Box) can be used to identify the effects of the selected firm-level, industry-level, and business cycle effects on the cost of equity (*COE*).[14] In this model, we explicitly control for firm size, as measured by the logarithm of book value of total assets (*BV*), financial leverage (*LEV*), effective corporate tax rates (TAX), the liquidity of a firm's stock (*LIQ*), and analysts' forecast dispersion (*DISP*). We control for changing economic conditions and industry effects by including year (YEAR) and industry (*IND*) dummy variables. The model also includes dummy variables denoting whether a firm is a U.S. firm (*US*) or a cross-listed Canadian firm (*XLIST*).

14. Again, we do not control for all possible influences on the cost of equity.

Box: Cost-of-Equity Regression

Using a panel data set, i.e., observations from many firms over many years, can present challenges for regression analysis, since the independent variables will vary both by time and by firm. This is complicated by the presence of time-invariant (dummy) variables. We overcome these difficulties by taking a two-step approach. In the first stage, a fixed-effects model is run using the time-varying independent variables:

$$COE_{i,t} = \alpha + \beta_{XLIST} XLIST_{i,t} + \sum_{t=1989}^{2006} \beta_{YEAR\_t} YEAR_{i,t} + \beta_{BV} BV_{i,t} + \beta_{LEV} LEV_{i,t} + \beta_{TAX} TAX_{i,t} + \beta_{LIQ} LIQ_{i,t} + \beta_{DISP} DISP_{i,t} + u_i + \varepsilon_{i,t}.$$

In the second stage, a weighted least-squares model is run, which regresses the firm fixed-effect coefficient (u_i) from the first-stage regression on the time-invariant independent variables (the U.S. and industry dummy variables), as well as the firm averages of the time-varying independent variables (to control for correlation between these variables and the firm fixed effects):

$$\hat{u}_t = \omega + \beta_{US} US_i + \sum_{k=1}^{K} \beta_{IND\_k} \overline{IND}_{i,k} + \gamma_{XLIST} \overline{XLIST}_i + \sum_{t=1989}^{2006} \gamma_{YEAR\_t} \overline{YEAR}_i + \gamma_{BV} \overline{BV}_i + \gamma_{LEV} \overline{LEV}_i + \gamma_{TAX} \overline{TAX}_i + \gamma_{LIQ} \overline{LIQ}_{i,t} + \gamma_{DISP} \overline{DISP}_{i,t} + \nu_i.$$

This set-up assumes common coefficients for all of the firms, both Canadian and U.S., in our sample and does not account for possible non-linear effects of our variables on the cost of equity.

With this approach, the resulting coefficient on the U.S. dummy variable (β_{US}) can be considered as the difference between Canadian firms' and U.S. firms' cost-of-equity financing (and, if multiplied by 100, it can then be expressed in basis points after accounting for the other regression variables).

The cost-of-equity differential between Canada and the United States over the 1988–2006 period is in the range of 30–50 bps.

Using a regression analysis that includes these firm characteristics, the results indicate that almost all of these control variables are statistically significant and have the expected relationship with the cost of equity (Table 2). For example, greater firm size is associated with a lower cost of equity; firms with more debt have a higher cost of equity; firms with higher stock liquidity have a lower cost of equity; and firms with more imprecise earnings estimates by analysts have a higher cost of equity. Once we account for all of these differences, plus the effects of industry membership and business cycles, U.S. firms in our sample appear to have a lower cost of equity, by approximately 47 bps, than do the Canadian firms. After subjecting our regression results to a number of sensitivity tests,[15] we conclude that for our sample of firms the cost-of-equity differential between Canada and the United States over the 1988–2006 period is in the range of 30 to 50 bps.

This analysis has improved upon previous studies by accounting for some of the differences across firms. It does not yet address, however, the possibility that differences in the risk-free rates faced by these firms could also be affecting their cost of equity. Failing to allow for different interest rate environments across countries may not lead to a fair comparison. The risk-free rate, typically represented by the longer-term sovereign bond yield, captures an important part of

15. All of our regression results are subjected to various robustness checks. In addition, results using other economic models are not significantly different from our own. See Witmer and Zorn (2007) for a discussion of these issues.

Table 2

Cost-of-Equity Regression Results, 1988–2006

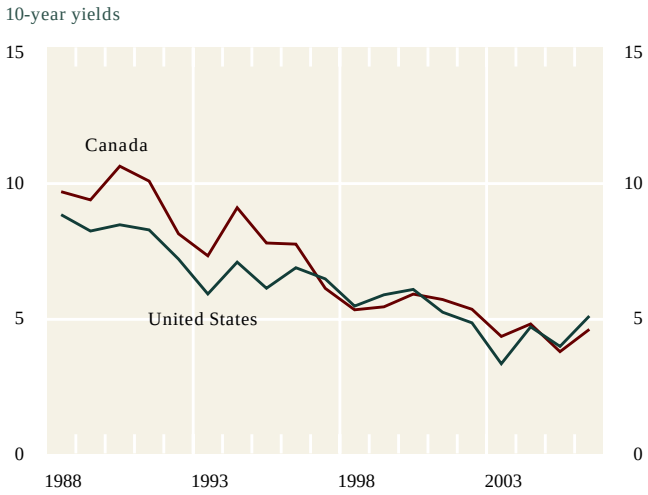
	β_i	*t*-statistic
Constant	12.015	26.21*
Size (total assets)	-0.247	3.87*
Financial leverage	0.64	12.43*
Taxes	-0.009	3.45*
Stock liquidity	-0.101	2.69*
Forecast dispersion	8.56	13.94*
U.S. firm	-0.465	3.40*

* Significant at 1 per cent

Notes: This table presents results for a 2-stage regression involving the U.S. dollar nominal cost of equity for Canadian and U.S. firms. For convenience, we do not report values for industry, year (business cycle), and cross-listed dummy variables. Absolute values of *t*-statistics are adjusted for heteroscedasticity of errors at a firm level.

Chart 3

10-Year Government Bond Yields, 1988–2006



the macroeconomic environment faced by firms. It reflects differences in monetary and fiscal policy regimes, including their effects on inflation uncertainty.

As Chart 3 shows, 10-year government bond yields declined between 1988 and 2006, roughly parallel with the decline in the cost of equity. However, there also appear to be two distinct interest rate periods. Canadian yields were much higher than U.S. yields during the first half of the sample period (1988–97), because investors demanded a higher risk premium to compensate for various factors, including high government debt levels and Quebec-related political uncertainty. Since 1997, there have been relatively small differences in yields between the two countries.

To examine the relation between bond yields and the cost of equity in our sample, we re-do our regression analysis in two different ways. First, we reformulate our regression equation to include nominal 10-year government bond yields as a right-hand-side variable[16] and find that a 100 bp increase in 10-year yields contributes to an increase of almost 20 bps in a firm's cost of equity.[17] With this specification, including the same regression variables plus 10-year yields, our tests are unable to conclude definitively that there is a difference between the Canadian and U.S. cost of equity. As a second test, we split our sample into two equal periods along the lines of the two interest rate periods that were identified: 1988–97 and 1998–2006. When our regression analysis is repeated, we find that, for 1988–97, the estimated differential between Canadian and U.S. cost of equity is very close to the full sample result in terms of sign, size, and statistical significance. However, in the latter period when sovereign bond yields were broadly similar in the two countries, the difference between the costs of equity in the two countries is lower, by about 20 bps, and is no longer statistically significant. This suggests that differences in longer-term sovereign bond yields may be a factor in explaining differences in the cost of equity.

Conclusions

The cost of equity for a firm is affected by several factors, some of which are related to characteristics of the firm itself, while others stem from the macroeconomic environment in which it operates. We find that greater firm size and greater liquidity of a firm's stock are associated with a lower cost of equity, while greater firm financial leverage and greater dispersion in analysts' earnings forecasts are associated with a higher cost of equity. Moreover, longer-term sovereign bond yields also seem to play a role in a firm's cost of equity. After taking firm-level and aggregate-level factors into account, the cost of equity in our sample was approximately 30–50 bps higher in Canada than in the United States over the 1988–2006 period. The cost-of-equity differential appears to be lower in the post-1997

16. In this model, the dependent variable is the firm's nominal cost of equity in its local currency.

17. Without year dummies, the estimated effect is closer to a 40 bp increase in a firm's cost of equity.

period, when sovereign bond yields were relatively similar in the two countries.

These results have policy implications. For example, since a smaller firm size adds to the financing cost of Canadian firms, promoting firm growth could have the positive effect of reducing the cost of equity. Higher forecast dispersion, or disagreement among equity analysts regarding firm earnings, is associated with a higher cost of equity. If better disclosure contributes to better forecasting of firm earnings, then improved disclosure regulation and practices in Canada might contribute to a lower cost of equity for firms. Perceived improvements to securities regulation and enforcement might also lead to greater trading and liquidity of Canadian stocks, in turn reducing the Canadian cost of equity. Finally, longer-term sovereign bond yields matter. This suggests that recent fiscal and monetary policy regimes, which have focused on pursuing a low debt-to-GDP ratio and anchoring inflation expectations to a low-inflation target, have had beneficial effects on the cost of capital for Canadian firms.

Longer-term sovereign bond yields seem to matter for a firm's cost of equity, suggesting that recent fiscal and monetary policy regimes have had beneficial effects for Canadian firms.

A sizable band of error accompanies the cost-of-equity estimates presented in this article, so a precise numerical value for Canadian cost of equity cannot be produced. In the same vein, Canada-U.S. differences are represented as an approximate value. To refine our estimates further, other methodologies could be applied and other factors could be considered, such as currency risk, inflation uncertainty, degree of market integration, personal taxes, and differences in regulatory environments. By incorporating proxies for these factors and perhaps extending our comparison to more countries, we might obtain better precision in the estimates and a broader international context for interpreting the results.

Literature Cited

Booth, L. 2001. "Equity Market Risk Premiums in the U.S. and Canada." *Canadian Investment Review* (Fall): 34–43.

Boritz, J. E. 2006. "Maintaining Quality Capital Markets through Quality Information." Discussion Paper Prepared for Capital Markets Leadership Task Force. Available at <http://www.cmltf.ca/index.cfm/ci_id/30808/la_id/1.htm>.

Claus, J. and J. Thomas. 2001. "Equity Premia as Low as Three Percent? Evidence from Analysts' Earnings Forecasts for Domestic and International Stock Markets." *Journal of Finance* 56 (5): 1629–66.

Hail, L. and C. Leuz. 2006. "International Differences in the Cost of Equity Capital: Do Legal Institutions and Securities Regulation Matter?" *Journal of Accounting Research* 44 (3): 485–531.

Hannah, B. 2000. "Approaches to Current Stock Market Valuations." *Bank of Canada Review* (Summer): 27–36.

Jorion, P. and W. N. Goetzmann. 2000. "A Century of Global Stock Markets." National Bureau of Economic Research (NBER) Working Paper No. 7565.

King, M. R. and E. Santor. 2007. "Family Values: Ownership Structure, Performance, and Capital Structure of Canadian Firms." Bank of Canada Working Paper No. 2007–40.

King, M. R. and D. Segal. 2003. "Valuation of Canadian- vs. U.S.-Listed Equity: Is There a Discount?" Bank of Canada Working Paper No. 2003–6.

Literature Cited (cont'd)

King, M. R. and D. Segal. 2006. "Market Segmentation and Equity Valuation: Comparing Canada and the United States." *Journal of International Financial Markets, Institutions & Money*. Forthcoming.

Task Force to Modernize Securities Legislation in Canada. 2006. *Canada Steps Up*.

Witmer, J. and L. Zorn. 2007. "Estimating and Comparing the Implied Cost of Equity for Canadian and U.S. Firms." Bank of Canada Working Paper No. 2007–48.

Bank of Canada Workshop on Derivatives Markets in Canada and Beyond

Toni Gravelle, Financial Markets Department

- *The development of derivatives markets supports both the efficiency and stability of the financial system.*
- *Because derivatives are designed to transfer risk rather than to transfer funds, the regulatory framework for derivatives can differ from that of stocks.*
- *Investors may not fully appreciate the risk involved in trading credit derivatives.*
- *The increasing reliance on the trading of risk-transfer instruments makes the financial system increasingly vulnerable to the possible evaporation of market liquidity.*

In September 2006, the Bank of Canada hosted a one-and-a-half-day workshop, *Derivatives Markets in Canada and Beyond*. The workshop focused on the prodigious and seemingly boundless growth in the volume and types of these risk-transfer instruments. It was also a forum where participants could exchange views on key developments in derivatives markets, voice concerns related to the risks associated with derivatives, and discuss areas where Canadian derivatives markets have led or lagged behind those in global financial centres. The event brought together market participants, regulators, policy-makers, and academics from various countries. This article presents the highlights of the workshop.

Background: Derivatives Basics

A financial derivative is an instrument whose payoff is typically linked to the underlying prices or value of interest rates or exchange rates, equity indexes, or other financial securities. More generally, the underlying price or payoff can be linked to almost anything, from the price of gasoline or wheat to the summer temperature readings in a particular city, or even the release of macroeconomic data, such as the size of the gross domestic product (GDP) or employment growth registered for the preceding quarter or month. Derivatives typically fall into one of the following categories: futures or forwards on equity, interest rates, and currency instruments; interest rate and currency swaps; options on equity, currency, interest rates, futures, and swaps; and interest rate caps, floors, and collars. As well, derivatives that are linked to the likelihood of default of one or several debt instruments have recently emerged and are one of the fastest-growing segments of this market.

The derivatives that trade on organized exchanges are futures and exchange-traded options. These derivatives consist of standardized contracts because exchanges are better suited to the trading of less complex and more "commoditized" financial instruments. A much broader and faster-growing range of derivatives instruments, including relatively more complex ones, are traded in the over-the-counter (OTC) markets by financial institutions, fund managers (including pension and hedge fund managers), and corporate treasurers. This segment of the derivatives market is also where innovation seems to flourish the most, with new, and at times complex, "made-to-order" derivatives contracts appearing regularly. Perhaps the most significant development in financial markets over the past five years or so has been the rapid development of credit derivatives. Discussions in several of the sessions focused or touched upon the evolution of credit derivatives.

In what follows, we provide a thematic synopsis of the various topics discussed during the workshop sessions.

Globalization and Technological Advances

There are two key drivers of innovation and growth in derivatives markets. The first is the globalization of finance, which has accompanied, and in many ways has been made possible by, the modernization and globalization of commercial and investment banking. The prodigious growth and development of derivatives markets are both symptoms and drivers of the globalization of finance. Specifically, derivatives markets have developed in parallel with the emergence of globally active financial intermediaries that handle the bulk of the international capital and capital flows in the major financial market centres, such as New York and London.

The prodigious growth and development of derivatives markets are both symptoms and drivers of the globalization of finance.

Although the trend to use globally active financial intermediaries has been evident in the banking industry since the 1980s and 1990s, several workshop participants noted the recent development of this trend within the pension fund sector in Canada. The sector has embraced active portfolio management, which largely entails the vigorous use of derivatives.[1] Moreover, Canadian fund managers have increasingly taken a global view of asset diversification and risk management, and larger Canadian funds have increasingly sought to create synthetic exposures to asset classes not readily available in Canada.[2] In doing so, they are more often seeking out larger foreign dealers to handle a growing share of their trading activity in derivatives markets.

These larger financial intermediaries tend to develop innovative derivatives structures in order to meet their own and their clients' needs and are better placed to take advantage of the economies of scale required to trade derivatives on a global basis. Workshop participants noted that an intrinsic characteristic of derivatives instruments is that they are designed to transfer risk, whereas stocks or bonds are designed to be an explicit claim on the stream of cash flows generated from the ownership of a financial asset in a certain jurisdiction. This characteristic also makes them more amenable to borderless trading, making OTC derivatives markets, in particular, more global in nature than, say, largely nationalistic equity markets.

The second driver is the rate of development of financial innovations and new derivatives instruments, which has been sustained by the continued advances in, and falling costs of, computing power and telecommunications.[3] Advances in information technology, coupled with financial institutions' drive to enhance returns and expand their global reach, have contributed to an environment in which financial intermediaries (and, to some extent, their clients, particularly hedge funds) are continuously introducing and / or embracing new derivatives instruments and advances in risk-management techniques. This in turn reinforces the increasing dependence of derivatives markets on technological advances for their development. Although there have

1. In addition, it was noted that several of the larger Canadian pension funds had become active in the New York and London credit derivatives markets.

2. For example, given the heavy weighting of the Toronto Stock Exchange index towards resource and financial-based stocks, Canadian pension fund managers have sought to increase their exposure to other corporate sectors via equity portfolio allocations in foreign jurisdictions, often using derivatives to take on the exposure or at least to hedge part of the foreign exchange exposure assumed when purchasing foreign stocks.

3. One workshop participant noted that the exponential growth in the volume of exchange-traded equity options traded in the United States was the result of two factors: increased competition resulting from technological innovations and electronic trading platforms, and regulatory changes aimed directly at inciting more competition across exchanges, such as allowing the cross-listing of equity options.

been numerous innovations in derivatives markets over recent years, none has been as important as the technological advances that have permitted the separation and active trading of credit risk. This is discussed in more detail below.

The Potential Benefits of the Growth of Derivatives

The first few sessions of the workshop described the trends in the growth of both OTC and exchange-traded derivatives in Canada and worldwide, with several participants noting that the evolution of derivatives markets has accelerated rapidly over the years, creating several potential benefits. A broader array of derivatives increases the ability of market participants to unbundle and separately trade the various risk components embodied in financial instruments.[4] This in turn allows market participants who trade derivatives to manage their financial risks more easily. The trading and transfer of risk also allows for the wider dispersion of risks across the financial system and increases cross-border capital flows. These factors have likely been key elements underpinning the greater resilience of financial institutions to market stresses over the years and have enabled markets to more effectively allocate capital to its highest return. Overall, developments in the derivatives market have contributed to more complete financial markets, and have improved market liquidity and increased the capacity of the financial system to effectively price and bear risk. The economy benefits as well, since broad, deep, and well-functioning capital markets contribute to a more efficient financial system, one which leads to stronger economic growth over time.

Developments in the derivatives market have contributed to more complete financial markets, and have improved market liquidity and increased the capacity of the financial system to effectively price and bear risk.

Perhaps the clearest evidence of the private benefits of derivatives is the continued spectacular growth of derivatives markets. As a result of the increasing demand for these products, the size of the OTC derivatives market reached a notional principal value of US$415 trillion by the end of 2006 (Bank for International Settlements 2007). Indeed, from 2005 to 2006, OTC derivatives markets grew by roughly 40 per cent, higher than the average annual growth rate for the previous four years (Chart 1).[5] At the same time, the size of the global exchange-traded derivatives market reached US$26 trillion in notional value by the end of 2006 (BIS 2007) (Chart 2). Turnover is similarly large. The most recent BIS data on OTC instruments and exchange-traded derivatives indicate that turnover rose from US$1.8 trillion in 2001 to roughly US$6.5 trillion per day in 2004, which converts to US$1,700 trillion on an annual basis. By comparison, nominal global GDP stood at US$51.5 trillion in 2006.

The sharp rise in OTC derivatives activity largely reflects the rapid growth of interest rate swaps and credit-default swaps. Workshop participants noted a similar trend in Canada. The Canadian interest rate swap (IRS) market has experienced exponential growth in volumes over the past five years (anecdotal evidence indicates growth of 25–50 per cent per year), accompanied by a significant narrowing of IRS bid / ask spreads. This growth has been driven mainly by the broadening of the Canadian IRS investor base to include foreign financial institutions and hedge fund as they seek to hedge their exposures to, or speculate on, cross-country differences in expected interest rate movements. The observed globalization of the investor base is also an important factor explaining the sharp rise in activity for Canadian exchange-traded financial derivatives. Between 2004 and 2006, the average daily volume of financial derivatives[6] contracts at the Montréal Exchange climbed by over 125 per cent. During this period, the proportion of foreign participants at the Montréal Exchange rose from approximately 40 per cent to close to 60 per cent.

4. It allows gasoline producers, for example, to separately measure and trade the price risk they face in selling gasoline from the risk they face in buying oil to produce the gasoline.

5. Note that the notional amounts overstate the risk embodied in the derivatives. The gross market value of derivatives, which measures the cost of replacing all existing contracts, represents a better measure of risk at any point in time. The gross amount at the end of 2006 was US$10 trillion, roughly the same amount as in 2005.

6. In discussing financial derivatives, we are explicitly excluding commodity futures contracts.

Chart 1

Volume of OTC Derivatives

Notional amount, US$ trillions

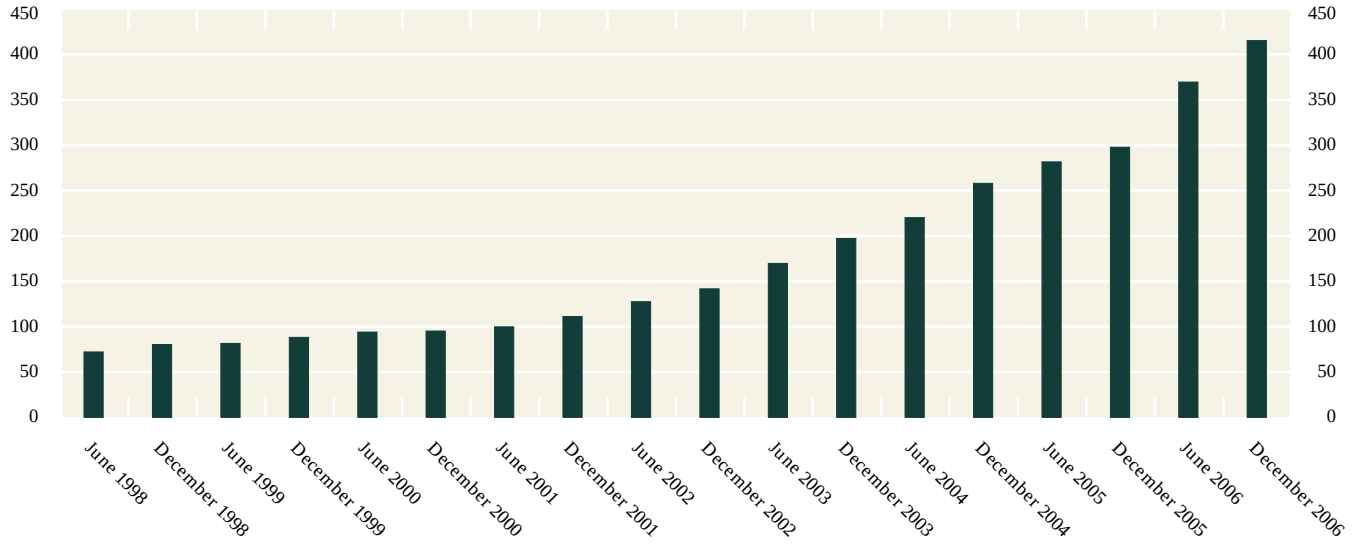
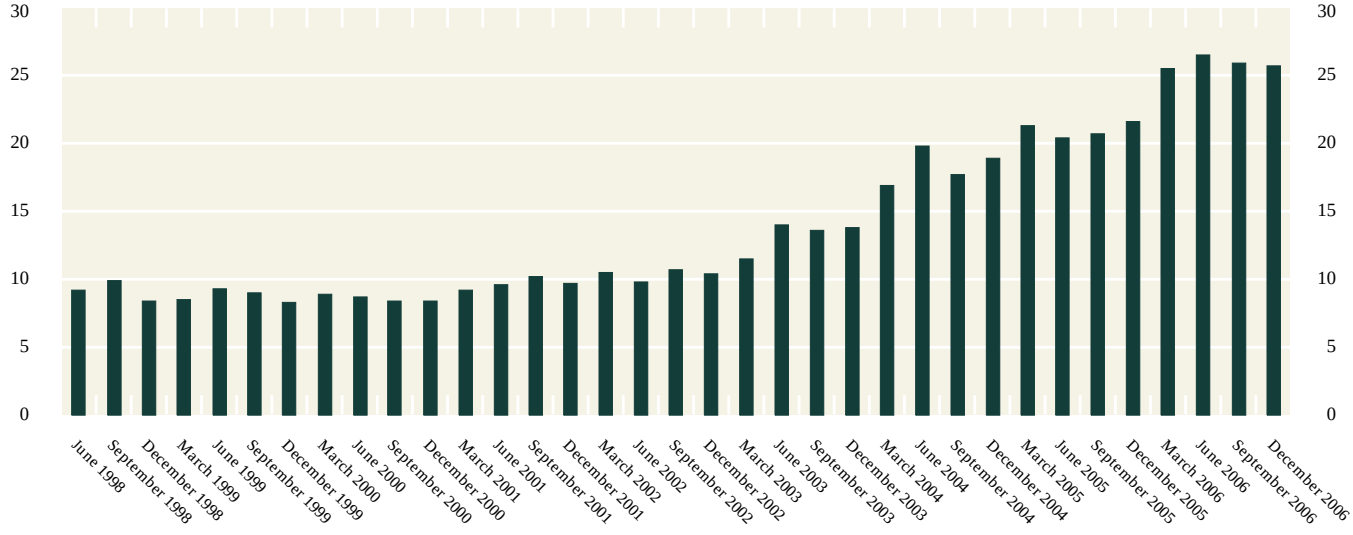


Chart 2

Volume of Exchange-Traded Derivatives

Notional amount, US$ trillions



Credit Derivatives: Growth and Challenges

Of all the segments of the derivatives market, credit derivatives received the most attention from workshop participants. The types of credit derivatives that have experienced the greatest growth are single- and multi-name credit default swaps (CDSs) and collateralized debt obligations (CDOs).[7] Growth in these instruments has been astounding. The notional amount of CDSs, for example, nearly doubled in each year for the past five years, with the amount outstanding growing from US$13 trillion in 2005 to nearly US$29 trillion in 2006. The amount of CDOs issued in 2006 was US$939 billion (BIS 2007), and the outstanding amount was estimated to be US$2.6 trillion.

Whether the benefits from the growth of derivatives markets and the associated innovations (described above) can be fully realized depends on how markets address the various financial-stability and risk-management issues posed by the use of these instruments. Three separate challenges related to credit and other derivatives were the focus of several discussions at the workshop.

The complexity of CDOs might also lead to the ultimate investors placing too great a reliance on the rating of the CDO tranches to guide their investment decisions.

The first challenge relates to the increasing complexity of these instruments, which, for many of the newer credit derivatives products, poses challenges to even the most sophisticated investors in terms of correctly modelling, understanding, and managing the embedded risk. It was noted that it is especially difficult to assess default correlations across several underlying reference assets in multi-name CDSs and in CDO tranches, making their valuation dependent on the underlying model's parameter assumptions about default correlations. Given this difficulty, concerns were raised about whether the ultimate holders of these instruments always fully grasp the nature of their risk exposures and how these exposures differ from those of more typical debt instruments, such as corporate bonds. The complexity of CDOs, as well as the requirement of many institutional investors to have their fixed-income holdings rated by a credit-rating agency, might also lead to the ultimate investors placing too great a reliance on the rating of the CDO tranches to guide their investment decisions.[8]

Secondly, there are concerns that secondary market liquidity for these instruments, particularly for CDOs, is less robust (or that these markets are more likely to become illiquid), owing to their complex model-driven valuation as well as to the lack of investor diversity and the concentration of intermediaries in these markets. Related to this, concerns were voiced that the cost of this potential market illiquidity was not fully reflected in the pricing of these instruments, leaving market participants exposed to sudden repricing and large mark-to-market losses in their portfolio holdings. This could trigger the simultaneous unwinding of crowded positions that would exacerbate the strains on market liquidity and could lead to detrimental knock-on effects on other debt markets and on financial intermediaries' balance sheets.[9]

Specifically, the advent of, and growth in, credit derivatives has essentially moved credit creation and the adjustment of credit exposures outside of the banking system. A sharp rise in asset-price volatility and concomitant drop in secondary market liquidity can now have a greater negative effect on credit creation than before. The greater connection between secondary market liquidity and the credit-creation mechanism is the necessary consequence of a system in which credit risk is "tradable" and dispersed outside the banking system, including among pension funds and leverage-investment vehicles such as hedge funds.[10]

7. CDSs basically provide insurance against the cost of default and various other credit events. That is, the protection buyer pays the protection seller periodic premiums in return for a payment if a credit event occurs. A CDO pools a portfolio of fixed-income assets into a tranched liability structure often seen in other securitized fixed-income instruments. The most common types of collateral for CDOs are asset-backed and corporate debt securities and syndicated loans. CDOs backed by loans are referred to as collateralized loan obligations (CLOs). Some do not in fact consider CDOs to be derivatives, but rather another type of fixed-income security. See Kiff and Morrow (2000), Kiff (2003), Reid (2005), and Armstrong and Kiff (2005) for more on credit derivatives in a Canadian context.

8. See International Monetary Fund (2006) for more on the possible over-reliance of institutional investors on credit-rating agencies.

9. For more on the market illiquidity issues, see Counterparty Risk Management Policy Group II (2005) and IMF (2006).

10. Since the workshop took place, these market-liquidity concerns related to credit derivatives have materialized as the events surrounding the global credit problems of August 2007 have unfolded. See Dodge (2007) and Longworth (2007) for details.

The issues posed by the infrastructure of OTC derivatives markets was the third challenge discussed at the workshop. The rapid growth of trading in credit and other OTC derivatives had (at the time of the workshop) largely outpaced the development of the infrastructure necessary to clear and settle those trades. Processing of completed trades was largely manual, and since trading volumes were increasing rapidly, derivatives dealers had accumulated a huge backlog of unconfirmed trades, even though they had greatly increased their back-office resources. Unconfirmed trades increase the potential for material mismeasurement and mismanagement of market and counterparty risk (see CRMPG II 2005; CPSS 2007). Steps to improve the situation were being undertaken jointly by regulators and the industry, but some workshop participants noted that, despite the substantial progress being made for uncomplicated derivatives, cleaning up the backlog for the more complex derivatives could still be challenging.

Inflation-Linked Derivatives

In recent years, the market for inflation-linked derivatives in Europe and the United States has grown rapidly. Futures contracts based on the U.S. consumer price index (CPI) and the euro zone harmonised index of consumer prices (HICP) (excluding tobacco) began trading on the Chicago Mercantile Exchange in 2004 and 2005, respectively. However, the largest segment of the inflation-linked derivatives market is the OTC inflation swap market that in essence began trading in 2001. An inflation swap is similar to standard interest rate swaps in which counterparties exchange cash flows based on a notional amount. For inflation swaps, counterparties exchange cash flows based on a fixed interest rate for variable payments linked to inflation.

Specifically, an inflation swap is a bilateral contractual agreement transacted in the OTC market. It requires one party to the contract (the inflation receiver) to make predetermined periodic fixed-rate payments in exchange for floating-rate payments linked to inflation from a second party (the inflation payer). Given that inflation swap contracts are traded OTC, a variety of contracts can be traded that incorporate different cash-flow structures to match the needs of the counterparties. The most popular type of contract, however, is the zero-coupon inflation swap, which has payments exchanged only on maturity.

In this contract, the fixed payments made by the inflation receiver for a *T*-year contract are calculated as follows:

$$fixed\ leg = (1 + fixed\ rate)^{T} \times notional\ value.$$

The variable inflation-leg payments made by the inflation payer are calculated as follows:

$$inflation\ leg = (inflation\ index_{at\ time\ T} / inflation\ index_{at\ inception}) \times notional\ value.$$

Although some swap contracts have extended further out, inflation swap maturities range, in general, from 1 year to 30 years.

The growth of this market resembles that of the IRS market in the early 1980s. It was noted during the workshop that the euro zone has the most liquid market, with an estimated total daily interdealer broker flow of roughly €500 million in the first part of 2006. In Europe and the United States, demand for inflation-linked swaps (i.e., demand to be the inflation receiver) stems from the demand by financial institutions and institutional investors to receive inflation-risk protection. In the United Kingdom and the United States, demand is mainly from pension funds, which seek to hedge long-term liabilities linked to inflation. In continental Europe, on the other hand, demand from financial institutions that sell inflation-protected instruments or inflation-linked deposits to retail or institutional investors is also significant.

The growth in inflation swaps activity has both coincided with, and been supported by, the significant increase in inflation-linked bond issuance in Europe and the United States (see Box). The same factors driving the demand for inflation swaps have also allowed for the greater issuance of inflation-linked bonds. However, inflation swap activity also relies on a sufficiently large and liquid inflation-linked bond market. That is, for dealers to make markets in inflation-linked OTC derivatives such as zero-coupon inflation swaps, they need to be able to economically hedge the inflation risk they take on as the inflation payer in one leg of the swap. Specifically, they must find an offsetting cash flow that is highly correlated with the cash flows that they are obliged to provide. Dealers have found that the most effective source of these offsetting cash flows is the purchase of government inflation-linked bonds denominated in the same currency (and based on the same inflation index) as the swap.[11] As such, large and liquid inflation-linked government bond markets with

11. If the dealer happened to be an inflation receiver in the contract, it would in this case hedge its position by selling inflation-linked bonds outright, if it owned them, or selling short, if it did not.

Box: Sovereign Inflation-Linked Bond Issuance

As of 2006, all G–7 countries* have inflation-linked bond issues, the value of which more than tripled between 2000 and 2006, reaching roughly US$1 trillion outstanding. The majority of the outstanding inflation-linked bonds are from the euro zone (largely France and Italy, which began issuing large volumes over this period), the United Kingdom, and the United States. These countries had an outstanding amount of US$260, $257, and $403 billion, respectively, at the end of 2006 (Hurd and Relleen 2006).

Although Canada has been issuing inflation-linked bonds (known as Real Return Bonds) since 1991, the volume (the outstanding amount of these bonds was $36 billion in 2006) has lagged substantially behind the volume of those issued in the euro zone, the United Kingdom, and the United States. Moreover, Canada has issued only four separate 30-year bonds over the years, while issuance in the three main inflation-linked bond jurisdictions has been across a variety of maturities, including 2-, 5-, 10-, 30-, and more recently, 50-year maturities.

* Sweden, Greece, and Australia also issue inflation-linked bonds.

a variety of outstanding bond maturities underpin the market-making activity of dealers in inflation swaps.[12]

Workshop participants noted that this likely explains the lack of inflation swap activity in Canada. It was pointed out that although Canadian institutional investors' demand for inflation-linked instruments had increased proportionate with the level of increase in the United Kingdom and the United States, the inflation-linked bond market does not have the required characteristics for dealers to effectively make markets in inflation swaps because of its insufficient size and liquidity.[13]

Another way dealers can hedge the inflation-payer obligations resulting from their inflation swap activity is to find investors who will engage in offsetting swap transactions rather than using inflation-linked bonds as a hedge. These investors or firms would tend to be those that have a "natural" source of inflation-linked cash flows. In the United Kingdom and elsewhere, corporations such as utilities, toll-road operators, or other infrastructure firms that have relatively stable inflation-linked revenues have increasingly become involved in the inflation swap market (McGrath and Windle 2006). These natural inflation payers have found participating in the inflation swap market an effective way to lower their cost of debt financing, given the robust demand from institutional investors for inflation protection. Workshop participants indicated that Canada could see a rise in inflation swap activity over time as natural Canadian inflation payers' awareness of the potential advantages of participating in the inflation-linked instruments increases.

Canada could see a rise in inflation swap activity over time as natural Canadian inflation payers' awareness of the potential advantages of participating in the inflation-linked instruments increases.

Concluding Remarks

Overall, the workshop discussions revealed how the recent rapid growth and development in derivatives markets are in many ways leading or reinforcing the

12. As an alternative to using inflation-linked bonds, dealers could, in principle, hedge their inflation risk exposure via recently introduced inflation futures contracts. But, given that inflation futures maturities don't extend much beyond one year, these are largely used to hedge only short-term inflation swaps.

13. Unlike the other G–7 countries, the Canadian government faces constraints on its ability to increase the size of any segment of its bond-issuance program because it has for several years experienced budgetary surpluses and in turn has been prudently reducing the amount of marketable debt outstanding.

trend towards the globalization of financial markets. The workshop also highlighted that derivatives instruments are intrinsically designed to transfer risk and to aid in price discovery, rather than to invest funds in an explicit claim on financial capital such as stocks and bonds. For this reason, the regulatory framework for derivatives exchanges (and their clearing and settlement organizations) can be quite different from that applied to equity and debt securities, as is the case in the United States. There, the statutory powers of the Commodity Futures Trading Commission (CFTC) acknowledge the implicit global nature of the futures exchange business. Moreover, the CFTC's regulatory framework is much more principles-based than, for example, the regulatory regime governing equity securities, since it reflects the main purpose of futures products, which is risk shifting and price discovery. It reflects as well the nature of the complex and continually evolving derivatives markets. A more principles-based regulatory approach is better suited for rapidly adapting to changing business structures, the introduction of new products, and market development.[14]

The development of the derivatives market was seen by workshop participants as providing broad economic benefits. By transferring and managing more risk in the capital markets, the banking system and the overall financial system might not only become more efficient, but also more resilient to shocks. Moreover, the development of derivatives markets will not only support economic and financial efficiency, but will also further contribute to improved financial stability.

The concerns raised about the use of derivatives are often related to their innovative features and complexity. As is the case whenever broad and rapid adoption of substantially new financial instruments occurs, there is the concern that market participants are not completely aware of, or do not fully understand, the explicit or implicit risks that arise in trading credit derivatives. History has shown that when this is the case, it often leads to an overextension of risk taking, a mispricing of financial instruments, and a hidden buildup of financial system vulnerabilities. Workshop discussion further highlighted how financial system distress is more likely to involve the evaporation of market liquidity in credit derivatives markets and to have far-reaching cross-border effects, given both the greater dependence of the credit-creation process on market liquidity (and in turn on an effective secondary market price-discovery process) and the globalization of finance. The events surrounding the August 2007 credit market strains would seem to bear out these concerns.

14. Note that the Autorité des marchés financiers in August 2007 published for comment a proposed framework for the regulation of derivatives markets in Québec that is based on core principles.

Literature Cited

Armstrong, J. and J. Kiff. 2005. "Understanding the Benefits and Risks of Synthetic Collateralized Debt Obligations." *Bank of Canada Financial System Review* (June): 53–61.

Autorité des marchés financiers. 2007. "Proposed Derivatives Framework." 10 August. Available at <http://www.lautorite.qc.ca/userfiles/File/projets-speciaux/produits-derives/Derives_consultation_ang_2007_08.pdf.

Bank for International Settlements (BIS). 2007. "Regular OTC Derivatives Market Statistics." May. Available at <http://www.bis.org/publ/otc_hy0705.htm>.

Committee on Payment and Settlement Systems (CPSS). 2007. "New Developments in Clearing and Settlement Arrangements for OTC Derivatives." Basel: Bank for International Settlements.

Counterparty Risk Management Policy Group II (CRMPG II). 2005. *Toward Greater Financial Stability: A Private Sector Perspective*. Report. 27 July. Available at <http://www.crmpolicygroup.org/index.html>.

Dodge, D. 2007. "Turbulence in Credit Markets: Causes, Effects, and Lessons To Be Learned." Speech to the Vancouver Board of Trade, Vancouver, British Columbia, 25 September.

Literature Cited (cont'd)

Hurd, M. and J. Relleen. 2006. "New Information from Inflation Swaps and Index-Linked Bonds." *Bank of England Quarterly Bulletin* 46 (1): 24–34.

International Monetary Fund (IMF). 2006. "The Influence of Credit Derivative and Structured Credit Markets on Financial Stability." In *Global Financial Stability Report*, 57–84. April. Available at < http://www.imf.org/External/Pubs/FT/GFSR/2006/01/pdf/chp2.pdf?>.

Kiff, J. 2003. "Recent Developments in Markets for Credit-Risk Transfer." *Bank of Canada Financial System Review* (June): 33–41.

Kiff, J. and R. Morrow. 2000. "Credit Derivatives." *Bank of Canada Review* (Autumn): 3–11.

Longworth, D. 2007. "Liquidity, Liquidity, Liquidity." Speech to the Investment Industry Association of Canada, Toronto, Ontario, 3 October.

McGrath, G. and R. Windle. 2006. "Recent Developments in Sterling Inflation-Linked Markets." *Bank of England Quarterly Bulletin* 46 (4): 386–96.

Reid, C. 2005. "Credit Default Swaps and the Canadian Context." *Bank of Canada Financial System Review* (June): 45–51.

The *Canadian Journey*: An Odyssey into the Complex World of Bank Note Production*

Jill Moxley, Helen Meubus, and Maura Brown, Communications Department

- *The Bank of Canada is responsible for supplying Canadians with bank notes that are secure and readily accepted by the public. Note security has always been the primary consideration behind the design and artwork of bank notes.*
- *To date, the Bank has developed a new series of bank notes approximately every 15 years as a way of staying ahead of counterfeiting.*
- *Rapid advances in computer technology that began in the late 1980s have changed the counterfeiting environment. In the process of designing and producing the* Canadian Journey *series of bank notes at the start of this decade, the Bank faced several challenges, including a sudden surge in counterfeiting activities.*
- *These challenges prompted the Bank to develop a new approach to its currency program. In 2002, the Bank adopted a comprehensive strategy to address counterfeiting: complementing increased bank note security with extensive public education and an intensified focus on law enforcement.*
- *This strategy, which involves active collaboration with the Bank's partners in the retail, law-enforcement, and legal communities, has reduced counterfeiting and increased public confidence in Canadian bank notes.*
- *The* Canadian Journey *series combines world-class security features and special user features with designs that represent the aspirations and accomplishments of Canadians.*

The creation of secure, attractive, and durable bank notes has always been challenging, but never more so than in the past decade, because of the rapidly improving technology available to counterfeiters. This article traces the development of the Bank's new anti-counterfeiting strategy as it evolved in parallel with the production of the new *Canadian Journey* series of notes.

To provide some background, the counterfeiting environment that evolved in the decades prior to the launch of the *Canadian Journey* series is described, together with details of the security features used in previous note series. This is followed by a discussion of the challenges encountered in developing the new series and the Bank's response. The final section reviews the valuable lessons that the Bank learned as it developed the *Canadian Journey* series.

Background: Counterfeiting and Confidence

Because bank note design has always been focused on ensuring that notes are accepted by the people who use them and that public confidence is maintained, security has been the main consideration in the design and artwork of bank notes (Bank of Canada 2006; Lefebvre-Manthorp 1988). Throughout the nineteenth and most of the twentieth century, along with the specialized paper on which bank notes are printed,

* Pierre Duguay, Charles Spencer, and Ianthi Vayid provided valuable guidance in the preparation of this article. Thanks are also due to Lea-Anne Solomonian, and to Micheline Lefebvre-Manthorp, Harry Hooper, and Robert Dolomont, who supplied useful background material.

the artwork itself was a major security feature. The ornate engraving and intaglio[1] created designs that made notes difficult to reproduce, thus protecting them from counterfeiting.[2]

Security has been the main consideration in the design and artwork of bank notes.

Since opening for business in 1935, the Bank has developed a new note design approximately every 15 years, because experience has shown that familiarity with a design allows counterfeiters to produce increasingly sophisticated copies that are accepted by the public. For example, it was not until 13 years after it was issued that the 1954 *Canadian Landscape* series was heavily counterfeited. In 1973, counterfeiting of the $50 note from the 1954 series pushed counterfeiting levels to over 950 counterfeits detected per million genuine $50 notes in circulation. For many years afterwards, retailers were reluctant to accept the $50 note. In 1973, the $5 note from the 1954 series was also heavily counterfeited.[3]

The Bank responded to these increased levels of counterfeiting with the multicoloured *Scenes of Canada* series, issued between 1969 and 1979. Because the range of tints beneath the dominant colours on these notes could not be easily reproduced through offset printing, the number of counterfeits fell dramatically (Chart 1).

Before the 1980s, counterfeiting was largely the domain of organized criminals, because the expert skills and expensive equipment required to produce counterfeits could only be financed by the extensive distribution networks maintained by such groups. The development of colour photocopiers in the 1980s, however, meant that counterfeits could be produced using the costly and sophisticated equipment available in some offices and copy shops. As the 1990s progressed, so did technological innovation. Before long, anyone with a personal computer and an inkjet printer could produce passable facsimiles of the images on bank notes.

With the 1986 *Birds of Canada* series, the Bank introduced a state-of-the-art security feature: the OSD, or optical security device, which anticipated the use of colour photocopiers for counterfeiting. A rectangular foil patch applied to the upper left-hand corner of the higher-denomination notes ($20 and up), the OSD shifted from gold to green when the note was tilted and was extremely difficult to reproduce.

Developed in partnership with the National Research Council, the OSD was first applied to the $50 note in December 1989. As a security feature, it was easy to check, and because its colour-shifting property could not be reproduced by colour copiers or scanners, it protected the high-denomination notes in the *Birds of Canada* series from counterfeiting for most of the 1990s.

Two other security features introduced with the *Birds of Canada* series were not apparent to the public. These were an anti-copier feature, which causes commercial photocopiers to recognize and refuse to copy a bank note, and a digital watermark, which has the same effect on most personal computer printers and scanners. However, it took some time for machines that could read these codes to become widely deployed.

Challenges in Producing a New Note Series

The development of the *Canadian Journey* series took place in this environment of rapidly accelerating technological progress. In 1997, the Bank established a currency development team and, the following year, began a design partnership with its two security printers: BA Banknote (now BA International Inc.) and the Canadian Bank Note Company, Limited.

The team faced several key issues in developing the new note series: an environment of budgetary restraint in the public sector; the search for a more secure substrate;[4] the choice of a design theme; the need to improve the readability of the notes by the visually impaired; and a surge in counterfeiting. In managing these chal-

1. Intaglio is a printing technique in which an image from a design is cut or etched into a metal plate, resulting in the image being reproduced in raised ink.

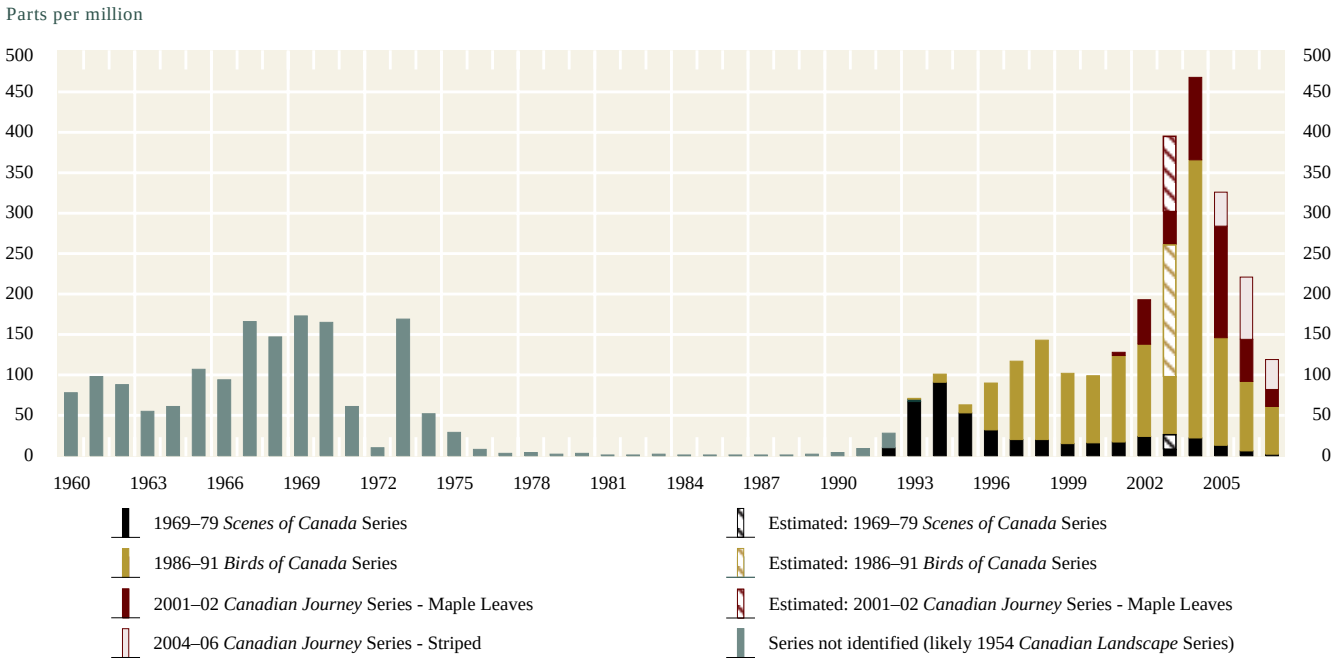
2. For an illustrated history of bank note design at the Bank of Canada, including details on the design and security features of previous note series, see Bank of Canada (2006).

3. Quarterly statistics on counterfeiting by series and by denomination are published in Table B4 of the monthly *Bank of Canada Banking and Financial Statistics*, now available at<http://www.bankofcanada.ca/pdf/bfs.pdf>.

4. Substrate is a term used in the security printing industry to refer to security paper that in itself offers basic security features such as fibre, a security thread, and a watermark. Substrates can also be made of polymer or consist of layers that sometimes sandwich an inner layer of plastic.

Chart 1

Counterfeit Canadian Bank Notes Passed for Every One Million Genuine Notes in Circulation (PPM)



lenges, the Bank learned several lessons, which it incorporated into a new anti-counterfeiting strategy, discussed later in this article.

The development of the Canadian Journey *series took place in this environment of rapidly accelerating technological progress.*

Budgetary restraint

In 1995, the federal government began a program review and a sweeping cost-cutting exercise to restore fiscal balance. The Bank embraced this policy direction and began its own review of operations in order to find more economical ways of conducting its business. In the process, the Bank closed seven agencies across the country in 1997–98 and outsourced the functions of note recirculation and inventory holding to the major financial institutions (Bilkes 1997).

The note-issuing department of the Bank, reacting to increased counterfeiting activity, had been advocating the development of a new series before obtaining approval to go ahead in 1997. But the emphasis on cost cutting in the mid-1990s, as well as the influence of a number of other factors, affected some of the Bank's choices as it moved forward with plans for the new series. Counterfeiting levels had been relatively low through to the mid-1990s, and the increasing popularity of debit cards suggested that there would be a reduction in the use of cash. Some financial observers were even musing about the possibility of a cashless society.[5] The Bank reasoned that if Canada was moving in that direction, perhaps it should be cautious about investing in costly research and development for bank notes. Having spent millions of dollars over 16 years to develop the OSD, the Bank concluded that it might be more cost effective to buy products developed by the bank note industry. It focused its efforts on producing a new series of economical notes that would not only be more secure, but would last longer in circulation and have a life-cycle cost per note the same as, or lower than, that of the *Birds of Canada* notes, thanks to new developments on the substrate front.

5. See, for example, Bank for International Settlements (1996). The Bank of Canada was also exploring the issue (Stuber 1996).

The search for a substrate

For the substrate on which the notes of the new series would be printed, the Bank investigated several options on the market or under development before focusing on an experimental new material being developed by a major Canadian paper manufacturer. Trademarked as Luminus, this new substrate was a sandwich of durable polymer laminated between two sheets of paper (McGovern 1995). The longer life expectancy of notes printed on Luminus would reduce the cost of replacing worn notes. The internal polymer layer was also thought to offer increased security because it could carry a coloured image similar to a watermark.

Between 1995 and 1998, Luminus was tested in active circulation with 100,000 *Birds of Canada* $5 notes. No major problems were identified, and in June 1998, the Bank was preparing to use Luminus as the substrate for the first two denominations of the *Canadian Journey* series—the $10 and $5 notes. In September 1999, this decision was extended to the higher denominations as well. However, technical issues with the production of Luminus, as well as questions about its market potential, led the owner of the technology to withdraw its offer to supply the product in December 1999.

As it became clear that Luminus would not be ready in time for the introduction of the first note in the *Canadian Journey* series (the $10 note), the Bank opted to issue the note on a 100 per cent cotton fibre substrate with surface characteristics similar to those of Luminus. This substrate would permit a seamless transition to Luminus when it became available.[6] In January 2001, the new $10 note was issued using this substrate, followed a year later by the $5 note on the same substrate.

Like the *Birds of Canada* notes, the original *Canadian Journey* notes featured microprinting, fine-line patterns, and intaglio print, as well as three new security devices:

- *A hidden number* to the left of the portrait that became visible when the note was tilted horizontally.
- *Fluorescent elements* consisting of: (i) random fibres on the face and back of the notes that glowed red or yellow under ultraviolet light; and (ii) printed images that appeared in blue over the portrait on the face of the notes under ultraviolet light. This relatively simple fluorescent feature was later developed into the more sophisticated fluorescence found on subsequent issues in the *Canadian Journey* series.
- *Three gold maple leaves* that were embedded in the design of the note face as an improvised substitute for the Luminus "watermark." Printed in a pale iridescent ink, the leaves reflected the light when the note was tilted.

These security features did not prove sufficient to protect the *Canadian Journey* $10 and $5 notes against counterfeiting.

Choosing a theme

A major challenge in the development of any new series is the choice of a theme for the series and then the selection of specific subject material for each denomination. In 1997, for the first time, the Bank held public consultations across the country on themes for the new notes. The resulting choice was images of Canadian wildlife. Prototypes were designed and presented for approval.[7]

Consultations generated public interest in the notes and greatly contributed to the public's ultimate acceptance of the new series.

However, with the series being launched at the beginning of the new millennium, the government proposed developing a theme that would project a more modern image of Canada, one that reflected the country's diversity, history, and values. Over the next year, new rounds of public consultations were held. Once themes for the individual notes were chosen and approved, the Bank also consulted interest groups that had a stake in the content and accuracy of the various note

6. The Bank obtained the Canadian rights to Luminus and continued to develop it with the help of other partners, since it remained convinced of the value of the substrate.

7. Under the Bank of Canada Act, S. 25(4), the "form and material" of bank notes are subject to approval by the Minister of Finance. These prototypes can be seen in *The Art and Design of Canadian Bank Notes* (Bank of Canada 2006, 100).

Box 1: *Canadian Journey* Design Themes

The design themes on the back of each note are:

$5 note, Children at Play. The images of children skating, tobogganing, and playing hockey represent young Canadians as the future of our nation, and play as a healthy part of their physical, social, and cultural development.

$10 note, Remembrance and Peacekeeping. The design commemorates the role of Canadians in past wars and evokes Canada's part in peacekeeping missions around the world.

$20 note, Arts and Culture. Internationally renowned artist Bill Reid (1920–98) drew inspiration from the Haida culture of Canada's northwest coast to create the works pictured: *The Spirit of Haida Gwaii, The Raven and the First Men, Haida Grizzly Bear,* and *Mythic Messengers.*

$50 note, Nation Building. To mark the accomplishments of women who campaigned for equal rights and social justice, the design features the statue of the Famous Five that can be seen on Parliament Hill and in Olympic Plaza in Calgary, Alberta, as well as the Thérèse Casgrain Volunteer Award medallion.

$100 note, Exploration and Innovation. A map of Canada drawn by Samuel de Champlain in 1632 and a birchbark canoe are paired with modern symbols of Canadian achievements in cartography and telecommunications.

designs. While consultations with these stakeholders[8] required additional time and work, they generated public interest in the notes and greatly contributed to the public's ultimate acceptance of the new series (Box 1).

Features for the blind and visually impaired

In addition to leading-edge security features, *Canadian Journey* notes incorporate new features for the blind and visually impaired. In the 1970s, consistent with the new Canadian Human Rights Act, the Bank began to review and assess possible ways of making bank notes more user-friendly for this group. Research indicated that a hand-held electronic note reader would be a better solution than such features as Braille,which many blind and visually impaired people cannot read, or different-sized denominations, which would add significantly to the costs of processing bank notes. The note reader, which could distinguish among the different denominations and read them out in a synthesized voice, was first distributed to users in 1990 by the Canadian National Institute for the Blind (CNIB) on behalf of the Bank.

In developing the *Canadian Journey* series, the Bank worked closely with the CNIB to explore and test alternative features. These consultations resulted in notes with many more user features that now serve a much wider group of individuals in the blind and visually impaired community:

- The colours on the notes were strengthened to assist people who have trouble distinguishing between browns and reds (the colours of the $100 and $50 notes) and purples and blues (the colours of the $10 and $5 notes).
- The numerals were enlarged and displayed on the front and back of the notes on two different, high-contrast backgrounds.
- A tactile feature was developed by the Canadian Bank Note Company in collaboration with Queen's University, the CNIB, and the Canadian Council of the Blind. While not Braille, the particular patterns of raised dots in the upper right-hand corner of each denomination identify the value of the note.
- The electronic note reader was made smaller, lighter, and more versatile.

The new features were very well received within the blind and visually impaired community, which numbers over 100,000. Upgrading the bank notes was a forward-looking initiative since, as the baby boom generation ages, the number of visually impaired users is expected to grow.

8. Stakeholders provided expert guidance on images to depict each theme. For example, stakeholders for the $100 note, with its theme of Exploration and Innovation, included Natural Resources Canada, the Canadian Space Agency, the National Archives of Canada, and the Canadian Canoe Museum.

Box 2: *Canadian Journey* Security Features

The key security features of the *Canadian Journey* series are:

Metallic holographic stripe—Along the length of the stripe, multicoloured numbers and maple leaves change colour through the various shades of the rainbow. Half of each maple leaf reflects one colour, half another colour. Small numbers that appear in the background of the stripe indicate the denomination of the note.

Ghost image (watermark)—A small, ghost-like image of the portrait and the denomination number appear on the front and back of the note when it is held to the light.

See-through number—Like the pieces of a jigsaw puzzle, irregular marks on the front and back of the notes (between the watermark and the large numeral) form a complete and perfectly aligned number representing the denomination when the note is held to the light.

Security thread—A solid vertical line can be seen from both sides when the note is held to the light. On the back of the note, the security thread is visible as a dashed line that changes colour from gold to green when the note is tilted. The thread contains the text "CAN" and the denomination number.

Escalation of counterfeiting

In 2001, highly deceptive *Birds of Canada* $100 notes featuring high-quality paper, fluorescent elements, and a copy of the OSD began turning up in stores along the length of the Windsor-Montréal corridor. In all, over 60,000 of these sophisticated counterfeit notes have been found in circulation. Some retailers in the concerned areas began displaying signs notifying their customers that they would no longer accept $100 notes. Fuelled by media stories, their refusal soon caused retailers across Canada to follow suit. At the peak, approximately 10 per cent of Canadian retailers were posting signs indicating their refusal to accept $100 notes.

As well, counterfeiting of the $20 *Birds of Canada* note increased markedly between 2002 and 2004, and counterfeiting of the new $10 *Canadian Journey* note rose dramatically. At the peak in 2004, counterfeiting of the $10 and $20 notes reached 1,292 and 601 parts per million, respectively.

A New Strategy

The Bank's response to this surge in counterfeiting was to adopt a comprehensive strategy designed to meet both the challenges posed by counterfeiters and the needs of the public. Under the new strategy, adopted in early 2002, the Bank would increase the intrinsic security of notes, build awareness of security features through education, promote the deterrence of counterfeiting by law-enforcement agencies and Crown prosecutors, and remove notes from older series from circulation.

The Bank's response was to adopt a comprehensive strategy designed to meet both the challenges posed by counterfeiters and the needs of the public.

Increased security

As it became clear that Luminus would not be ready on time to meet the immediate need to quickly issue higher-security notes, the bank note team began to investigate proven substrates for the higher-denomination notes in the *Canadian Journey* series ($20, $50, and $100).[9] The Bank chose a cylinder-mould-made substrate of 100 per cent cotton fibre that could carry the sophisticated, but user-friendly, new security features: a metallic holographic stripe, a watermark, a see-through number, and a colour-shifting windowed thread developed from the Bank's own OSD material (Box 2). This substrate was widely used by central banks in Europe and elsewhere.

9. At that time, Luminus was still the substrate of choice to reissue the lower-denomination notes.

Once government approval for this substrate was obtained, a European security-paper manufacturer was selected as the supplier through a competitive international tender, and all three *Canadian Journey* notes were issued in quick succession during 2004.[10,11] Luminus did not measure up to the mould-made watermark, so the $10 and $5 notes were later upgraded with the more advanced features of the other denominations in the series and were reissued in 2005 and 2006, respectively.

At the same time, the Bank worked with financial institutions to accelerate the removal of older series notes from circulation.

Building awareness

In 2004, the Bank undertook a national communications campaign to increase awareness of the security features used on the $20, $50, and $100 notes among retailers, financial institutions, and law-enforcement agencies. The campaign involved working with the media and providing training and information sessions conducted by Bank representatives at the regional offices (Bank of Canada 2005, 30).

In addition, with the issue of the high-denomination *Canadian Journey* notes in 2004, for the first time in the Bank's history, the unveiling of each new note took place several weeks or months prior to its release. These launches, to which the media and the Bank's partners were invited, occurred in the major cities across Canada where the Bank has regional offices. The advance publication of information about the notes gave the manufacturers of bank-note-processing equipment time to make adjustments and allowed retailers to familiarize themselves with the new notes so that they would readily accept them when they were put into circulation.

Promoting counterfeit deterrence

Education had been part of the Bank's currency program since the opening of regional offices in 1997, but playing an active role in counterfeit deterrence was new. In meetings with the law-enforcement and legal communities, the Bank began to draw attention to the social and economic costs of counterfeiting, which go far beyond the monetary loss suffered by the person who receives a worthless piece of paper (Healy 2002; Chant 2004). The Bank devised tools which served as resources for retailers, police, and public prosecutors in deterring counterfeiting activities.[12] Partnering with the Bank in counterfeit deterrence led the Royal Canadian Mounted Police (RCMP) to identify economic integrity as one of its strategic priorities in 2005. More recently, in May 2007, the Bank, the RCMP, and the Minister of Public Safety announced the launch of nationwide Integrated Counterfeit Enforcement Teams (ICETs) to fight organized counterfeiting.[13]

As a result of this new strategy, the number of counterfeit notes detected in circulation between 2004 and 2006 fell by almost half, and of these, nearly three-quarters were notes from older series (Bank of Canada 2007, 24). Improved public confidence in bank notes is evident from the decline in the percentage of stores that display signs refusing $100 notes, from a peak of 10 per cent to between 2 and 4 per cent.

Lessons Learned

Chief among the lessons that the Bank learned is the importance of having a comprehensive strategy in which all of the elements work together to reinforce public confidence in Canada's bank notes. Conducting research into and employing the most advanced security features creates secure bank notes, while promoting awareness of the security features supports their proper recognition and use. Removing from circulation notes from older series also reduces confusion among the public and retailers. As well, sensitizing the law-enforcement and legal communities to the costs of counterfeiting helps them to reinforce the message that counterfeiting is a serious crime. A number of specific lessons contributed to the development of this strategy.

Security is the key priority

Designing and producing secure bank notes has always been a primary objective of the Bank's currency program. The experiences with the *Canadian Journey* series reinforced the necessity of adhering to this goal and showed that low-denomination notes are now as vulnerable to counterfeiting as higher-denomination notes, and require the same level of protection.

10. Substrates offered by domestic suppliers could not provide the desired level of security.

11. The $20 *Canadian Journey* note won the "Bank Note of the Year for 2005" award from the International Bank Note Society for "its well-balanced design, strong images, and advanced security features" (IBNS 2007).

12. Some of this material is available on the Bank of Canada's website at <http://www.bankofcanada.ca/en/banknotes/counterfeit/index.html>.

13. The ICETs are based in Toronto, Montréal, and Vancouver, with counterfeit specialists also based in Halifax and Calgary.

Recognizing the value of partnerships in its efforts to combat counterfeiting, the Bank has intensified collaboration with its partners in the security printing industry, in equipment manufacturing, and at other central banks. It is a member of several international currency organizations, including the Central Bank Counterfeiting Deterrence Group, which shares information and collaborates in the development of solutions to counterfeiting at an international level.

Research and development must be continuous

In light of the accelerating pace of technological innovation and subsequent counterfeiting threats, as well as the long lead times involved in developing a new note series, ongoing research and development that is focused on preventive, high-quality security features is essential. Experience with the *Canadian Journey* series showed that it is easier to prevent problems than it is to fix them afterwards. With its new currency strategy, the Bank has adopted a more proactive approach to dealing with the ongoing threat of counterfeiting. The Bank has strengthened its research and development function to include surveying and monitoring activities in such areas as confidence levels and awareness of security features.[14] It is thus much better positioned to detect—and assess—counterfeiting threats as early as possible.

Research and development takes time, however, and is intrinsically risky. The Bank continues to develop the technology necessary to provide bank notes with the highest possible security, particularly where these requirements are not met by products currently on the market. But, when appropriate, it also purchases proven, leading-edge technology, which, in turn, requires continuous research into what is currently available. Such research also ensures that the Bank will always have a contingency plan in place to deal with unexpected surges in counterfeiting, which can erupt suddenly.

Plans for the next generation of bank notes call for a series that will be even more difficult to reproduce.

In the future, the Bank will likely increase the frequency with which it issues new series of bank notes. The Bank is currently investigating new materials and technologies that might be used to improve security, as well as the factors that facilitate verification by the public. Plans for the next generation of bank notes, to be launched in 2011, call for a series that will be even more difficult to reproduce and easier to authenticate.

Retailers and the public must be familiar with security features

As noted above, the Bank regularly conducts surveys that track public awareness of security features. As part of its ongoing initiatives to build awareness and promote counterfeit deterrence, the Bank will continue to familiarize retailers and the public with the security features and designs of any new notes before they are released into circulation.

Old notes must be removed from circulation as quickly as possible

Since notes from previous issues with less-secure features are more vulnerable to counterfeiting, their speedy removal from circulation is essential, but even this is not enough. In 2002 and 2003, there was a spate of counterfeit notes from a series that had long dropped out of circulation (see Chart 1). It is thus important to discourage the use of older notes in trade (Bank of Canada 2004, 28).

The demand for bank notes continues to grow

Although electronic payment methods are gaining in popularity, the volume of bank notes in circulation continues to grow in line with the economy. This indicates that bank notes are still a significant means of payment and store of value in the Canadian economy, and that the positive attributes of cash, namely, convenience, the protection of privacy, broad acceptance, and public confidence, continue to be valued (Taylor 2006). The Bank therefore needs to base its currency strategy on the assumption that bank notes will be part of the retail payment environment for the foreseeable future.

Conclusion

In dealing with the challenges of the *Canadian Journey* series, the Bank has developed a currency strategy designed to meet the requirements of a modern, complex economy. The goal of the strategy is, as always, to supply Canadians with high-quality bank notes that are readily accepted and secure against counterfeiting.

14. More details on the Bank's program of currency research are available at <http://www.bankofcanada.ca/en/fellowship/highlights_res_07.html#6>.

Literature Cited

Bank for International Settlements. 1996. *Implications for Central Banks of the Development of Electronic Money*. Basel: Bank for International Settlements.

Bank of Canada. 2005. *Annual Report, 2004*. Ottawa: Bank of Canada.

———. 2006. *The Art and Design of Canadian Bank Notes*. Ottawa: Bank of Canada.

———. 2007. *Annual Report, 2006*. Ottawa: Bank of Canada.

Bilkes, G. 1997. "The New Bank Note Distribution System." *Bank of Canada Review* (Summer): 41–54.

Chant, J. F. 2004. "The Canadian Experience with Counterfeiting." *Bank of Canada Review* (Summer): 41–54.

Healy, P. 2002. "Counterfeit Canadian Bank Notes." Available at <http://www.bankofcanada.ca/en/banknotes/pdf/healy_counterfeiting%20paper.pdf>.

International Bank Note Society (IBNS). 2007. "Bank Note of the Year for 2005." Available at <http://www.ibns.it/>.

Lefebvre-Manthorp, M. 1988. "Designing Canada's Bank Notes." *Bank of Canada Review* (January): 3–12.

McGovern, S. 1995. "Tough Luck, Counterfeiters." *Montreal Gazette*, 2 June.

Stuber, G. 1996. "The Electronic Purse: An Overview of Recent Developments and Policy Issues." Bank of Canada Technical Report No. 74.

Taylor, V. 2006. "Trends in Retail Payments and Insights from Public Survey Results." *Bank of Canada Review* (Spring): 25–36.

Bank of Canada Publications

Annual Report (published annually)*

Monetary Policy Report (published semi-annually)*

Monetary Policy Report Update (published in January and July)*

Financial System Review (published in June and December)*

Bank of Canada Review (published quarterly, see page 2 for subscription information)*

Speeches and Statements by the Governor*

Bank of Canada Banking and Financial Statistics*

Weekly Financial Statistics (published each Friday, available by mail through subscription)*

Renewal of the Inflation-Control Target: Background Information*

The Thiessen Lectures*

The Art and Design of Canadian Bank Notes
A journey behind the scenes to explore the demanding world of bank note design (published in 2006, available at Can$25 plus shipping costs).*

The Bank of Canada: An Illustrated History
Published in 2005 to celebrate the Bank's 70th anniversary, this book depicts the history of the Bank from 1935 until now (available at Can$25 plus shipping costs).*

A History of the Canadian Dollar
James Powell (2nd edition published December 2005, available at Can$8 plus GST and PST, where applicable)

The Transmission of Monetary Policy in Canada
(published in 1996, available at Can$20 plus GST and PST, where applicable)*

Bilingualism at the Bank of Canada (published annually)*

Bank of Canada Publications Catalogue, 2006*
A collection of short abstracts of articles and research papers published in 2006. Includes a listing of work by Bank economists published in outside journals and proceedings.

Planning an Evolution: The Story of the Canadian Payments Association, 1980–2002
James F. Dingle (published June 2003)*

About the Bank (published March 2004; revised edition, 2007)*

Conference Proceedings

Conference volumes published up to and including April 2005 are available on the Bank's website. Print copies can be purchased for Can$15 plus GST and PST, where applicable. Papers and proceedings from Bank of Canada conferences, seminars, and workshops held after April 2005 are now published exclusively on the Bank's website at www.bankofcanada.ca.

Technical Reports and Working Papers

Technical Reports and Working Papers are usually published in the original language only, with an abstract in both official languages. Single copies may be obtained without charge. Technical Reports dating back to 1982 are available on the Bank's website, as are Working Papers back to 1994. Consult the April 1988 issue of the *Bank of Canada Review* for a list of Technical Reports and Staff Research Studies published prior to 1982.

For further information including subscription prices contact:

Publications Distribution,
Communications Department,
Bank of Canada, Ottawa, K1A 0G9
Telephone: 613 782-8248

* These publications are available on the Bank's website, **www.bankofcanada.ca**