Do banks overstate their Value-at-Risk?

Christophe Péignon a,∗, Zi Yin Deng b, Zhi Jun Wang b

a HEC School of Management – Paris, 1 rue de la Liberation, 78351 Jouy-en-Josas, France
b Faculty of Business Administration, Simon Fraser University, 8888 University Drive, Burnaby, BC, Canada V5A 1S6

Received 14 September 2006; accepted 30 May 2007
Available online 4 September 2007

Abstract

This paper is the first empirical study of banks’ risk management systems based on non-anonymous daily Value-at-Risk (VaR) and profit-and-loss data. Using actual data from the six largest Canadian commercial banks, we uncover evidence that banks exhibit a systematic excess of conservatism in their VaR estimates. The data used in this paper have been extracted from the banks’ annual reports using an innovative Matlab-based data extraction method. Out of the 7354 trading days analyzed in this study, there are only two exceptions, i.e. days when the actual loss exceeds the disclosed VaR, whereas the expected number of exceptions with a 99% VaR is 74. For each sample bank, we extract from historical VaRs a risk-overstatement coefficient, ranging between 19 and 79%. We attribute VaR overstatement to several factors, including extreme cautiousness and underestimation of diversification effects when aggregating VaRs across business lines and/or risk categories. We also discuss the economic and social cost of reporting inflated VaRs.

© 2007 Elsevier B.V. All rights reserved.

JEL classification: G21; G28; G32

Keywords: Value-at-Risk (VaR); Capital requirement; Backtesting

1. Introduction

Current CEOs of commercial banks only have to worry about the remaining 1% of their profit-and-loss (P&L) distribution since nowadays most banks routinely compute daily 99% Value-at-Risk (VaR). The latter market risk measure summarizes the expected maximum loss over a target horizon (e.g. 1 day, 1 week) at a given confidence level (e.g. 95%, 99%). Despite vigorous criticisms over its mathematical properties (Artzner et al., 1999) and over its potential destabilizing effects on the economy (Danielsen et al., 2001; Basak and Shapiro, 2001; Leippold et al., 2006), VaR remains the reference measure for market risk. 1

In accordance with international standards, domestic regulators require all banks active in their jurisdiction to compute their VaR forecasts. In the US, public market risk disclosures are required under Financial Reporting Release Number 48 published by the SEC in 1997. In Canada, VaR calculation has been made compulsory since 1997 by the Office of the Superintendent of Financial Institutions Canada (OSFI) and VaR figures have been publicly disclosed since the late nineties. VaR disclosure accomplishes three main objectives. Firstly, it provides an all-inclusive measure of the bank’s market risk and its exposure to any kind of market risk. Secondly, it allows banks to assess their risk and to monitor closely any potential deviation from their targets. Finally, it induces “a sense of common purpose”, a sense of mutual responsibility toward the public, by imposing a certain discipline on banks. From a social perspective, VaR disclosure is supposed to bring about a more transparent, well-informed and prudent banking system. However, the empirical evidence presented in this study indicates that banks, whether American or Canadian, may overstate their VaR estimates.

∗ Corresponding author. Tel.: +33 139 67 94 11; fax: +33 139 67 70 85.
E-mail address: perignon@hec.fr (C. Péignon).

0378-4266/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.jbankfin.2007.05.014

“VaR gets me to 95% confidence. I pay my Risk Managers good salaries to look after the remaining 5%”.
Dennis Weatherstone, former CEO, J.P. Morgan

1 The view that common VaR-based risk modeling framework may destabilize financial markets was recently challenged by two empirical studies. Jorion (2007) shows that quarterly trading revenues are moderately correlated across US banks and Berkowitz and O'Brien (2007) find significant heterogeneity in US banks’ exposures to market factors. These findings do not corroborate systemic risk concerns.
of the amount of market risk borne by a given bank, which is supposed to reduce the information asymmetry between the firm and the market.\textsuperscript{2} Secondly, VaR estimates are translated into a capital charge that aims to provide a sufficient cushion for cumulative losses arising from adverse market conditions.\textsuperscript{3} Thirdly, it allows the regulator to assess the validity of the bank’s VaR model, a process known as backtesting.\textsuperscript{4} The basic idea is that an accurate VaR model, regardless of the statistical method it is based upon, should not lead to too many exceptions, i.e., days when the actual loss exceeds the disclosed VaR. For example, if a bank calculates its VaR assuming a one-day holding period and a 99% confidence level, then it is to be expected that, on average, trading losses will exceed the VaR measure on one occasion in one hundred trading days.

Most national supervisory authorities allow banks to use proprietary in-house models (internal models) for measuring market risk, as an alternative to a standardized measurement framework based on the main sources of market risk (Basel Committee on Banking Supervision, 1996a; Hendricks and Hirtle, 1997).\textsuperscript{5} A given internal model will only be accepted by the regulator if the bank can demonstrate the model accuracy through external validation. However, there is no obligation – nor any incentive – for the banks to disclose the specificities of their internal VaR estimation engines. Furthermore, VaR forecasts have direct implication in setting capital requirements for banks. Indeed, the capital charge for market risk for a bank using an internal model is to be the higher of either the VaR of the previous day or three times the average of the daily VaR of the preceding 60 business days. A penalty component shall be added to the multiplicative factor of three if the number of exceptions is abnormally large (Basel Committee on Banking Supervision, 1996b). For instance, based on a sample of 250 observations, a penalty ranging between 0.4 and 1 is imposed if the actual number of exceptions exceeds five.

Many market commentators have indicated that the high degree of autonomy granted to commercial banks in setting capital charges might have some perverse effects. In particular, banks may be inclined to underestimate their VaR in order to reduce their market risk charge. This intuitive view is also supported by formal theoretical models showing that it can be rational for a bank to intentionally understate its risk estimate. For instance, Lucas (2001, p. 826) concludes that “under the current regulatory framework, banks are prone to underreporting their true market risk”. Conversely, in their theoretical analysis of VaR-based capital requirements, Cuoco and Liu (2006, p. 362) find that “VaR-based capital requirements can be very effective in curbing portfolio risk but also in inducing revelation of this risk”. However, the empirical validity of these theoretical predictions still needs to be assessed.

To date, little is known about the actual accuracy of banks’ internal risk management models. An early unpublished manuscript by Cassidy and Gizecki (1997) reveals vast differences in the accuracy of five individual trading portfolios’ VaRs computed by an unidentified Australian bank between 1992 and 1995. Berkowitz and O’Brien (2002) published the first direct evidence on the performance of US banks’ internal VaR models. They show that aggregate VaR estimates are conservative and that they do not outperform forecasts based on simple econometric models, such as a GARCH model applied to the bank’s P&Ls. Using a sample of international banks, Pérignon and Smith (2007) report that VaR computed using historical simulation contains little information about future trading revenue volatility. Berkowitz et al. (2006) use daily VaR and P&L data generated by four separate business lines with a large international commercial bank. They find that the actual number of exceptions in each business line is either equal to or smaller than the expected number of exceptions.

Our paper contributes to the debate on the accuracy of the VaR models used by commercial banks by asking a simple question: Do banks overstate their VaR? We answer this question by studying actual VaR and P&L data from a sample of Canadian commercial banks. We find overwhelming evidence supporting the idea that commercial banks do exhibit a systematic excess of conservatism when setting their VaR. Out of the 7354 trading days analyzed in this study, there are only two exceptions whereas the expected number of exceptions with a 99% VaR is 74. This empirical result contradicts the common wisdom that banks intentionally underestimate their market risk to reduce their market risk capital charges. For each sample bank, we estimate from historical VaRs a risk-overstatement coefficient, which ranges between 19 and 79%. We identify two main causes of VaR overstatement. Firstly, banks do not want to gamble their reputation and aim to minimize the likelihood of having to disclose many exceptions. Indeed, the market is likely to severely punish any bank that is not able to correctly assess its exposure to market risk. Secondly, VaR overestimation arises when banks aggregate VaR estimates across business lines and/or risk categories without properly accounting for diversification effects. We also discuss the economic and social cost of reporting inflated VaRs.

Actual daily VaR and P&L are usually not publicly available to researchers in a machine-readable format. Indeed, daily VaRs are not private information since they are sometimes disclosed as a graph in the annual reports,
but they are not really public information either. We refer to this type of financial information as “seemingly public”. In response to this lack of ready-to-use data, we developed a data extraction technique allowing us to extract the data from the graph included in the banks’ annual reports. A direct implication of this data collection technique is that the daily VaR and P&L data used in this study are not anonymous. In particular, unlike all previous empirical studies, we do not scale banks’ data by their sample standard deviation to protect the confidentiality of individual institutions. The fact that our sample banks are not anonymous is useful because it allows us to provide full details about the banks’ balance sheets, regulatory capital, and VaR models. It also allows us to link the degree of VaR overstatement of a given bank to several key characteristics, such as the bank’s total derivatives position. Furthermore, our methodology to measure VaR overstatement at bank level and to assess the cost of VaR overstatement in terms of regulatory capital requires the use of unscaled data.

The rest of the paper is organized as follows. In the next section, we present an empirical analysis of actual VaR and P&L data for a sample of Canadian banks. For each bank, we backtest actual VaRs, compare banks’ VaRs with simple VaR estimates based on historical simulation and GARCH modeling, measure the level of risk overstatement, and assess the cost of over-reporting VaR. We also discuss several mechanisms that may lead a commercial bank to inflate its market risk disclosures. We summarize and conclude our study in Section 3.

2. Empirical analysis

2.1. Data

We study actual VaRs and P&Ls of six Canadian commercial banks over the period November 1, 1999–October 31, 2005. The six banks are Bank of Montreal (BMO), Bank of Nova Scotia or Scotiabank (BNS), Canadian Imperial Bank of Commerce (CIBC), National Bank of Canada (NBC), Royal Bank of Canada (RBC), and Toronto-Dominion Bank (TD). We study Canadian commercial banks because, unlike US commercial banks, they systematically display in their annual reports a graph of the daily VaR and P&L data.

The P&L variable measures the daily gain or loss on banks’ trading portfolios. To be consistent with VaR forecasts that rely on the assumption of no change in the portfolio during the holding period, P&Ls measure hypothetical changes in portfolio value that would occur if end-of-day positions were to remain unchanged. In our sample, all banks compute hypothetical P&Ls, except Bank of Nova Scotia and Toronto-Dominion Bank that report actual P&Ls. Only one sample bank, Bank of Montreal, acknowledges including certain fees and commissions directly related to trading activities, which may create some distortions in backtesting. This contrasts with the sample used in Berkowitz and O’Brien (2002, p. 1095), in which all sample banks include fee incomes in reported trading P&Ls.

We collect daily VaR and P&L data from the banks’ annual reports. For each bank/year, we retrieve actual VaRs and P&Ls using an innovative Matlab-based application that allows us to convert the graph from the annual report into a time series of daily values (see the Appendix A for details about the data extraction procedure). An implication of our approach is that the identity of the banks is preserved, which contrasts with previous studies relying on anonymous data that are scaled by their sample standard deviation prior to the analysis. A potential limitation of our approach is that the empirical analysis is based on estimated VaRs and P&Ls. In order to detect any discrepancy between estimated and actual values, we superimpose, for each bank and each year, the original graph from the annual report and the graph based on the extracted data. For all sample banks, actual and extracted data (i.e. VaRs and P&Ls) are not visually distinguishable. Furthermore, we gauge the accuracy of our data extraction method using simulations. We randomly generate 10 time series of 250 hypothetical daily, normally distributed VaRs, with a mean of $13 million and a standard deviation of $5 million that match average values observed in the real data. We upload the graph corresponding to each time series in our data extraction application and retrieve the underlying time series. Comparing the actual and estimated time series by means of a mean absolute error (MAE) leads to an average MAE of 0.91% with a standard deviation of 0.41%.

Based on our systematic visual comparison and our simulations, we are confident that the data used in this paper are reliable.

---

6 Actual sample periods vary across banks depending on data availability (see Table 1 for details).
7 US commercial banks very rarely report daily VaRs and P&Ls — two exceptions are Bank of America and Wachovia. They sometimes display a histogram of the daily P&L, which prevents applying our methodology to all large US banks. Over our sample period, the Canadian stock market experienced tremendous ups and downs: +67% in the S&P/TSX Composite Price Index between January 1999 and August 2000, −45% between September 2000 and September 2002, and +68% between October 2002 and October 2005. Furthermore, skyrocketing oil price and the strong appreciation of the Canadian dollar against the US dollar strongly affected the Canadian economy over our sample period.

8 Another limitation of our method of extracting data from graphs is that graphs might be smoothed transformations of raw input data. By only observing the graphs one cannot judge to what extent such smoothing might have taken place. This potential measurement error is beyond our control though. We thank a referee for pointing this out.
9 In our simulations, we do not allow for post-extraction adjustments. Conversely, in the empirical study, we superimpose the actual and extracted VaR time series and we keep iterating until we get a perfect match between the two lines. As a result, the extraction error is likely to be smaller in the empirical analysis than in the simulations. See Péron and Smith (2007) for further simulations and controlled experiments using our data extraction technique.
We present some descriptive statistics about the sample banks in Table 1. The six sample banks are the largest commercial banks in Canada and have the most important derivative positions. The term “Big Six” is frequently used to refer to the six biggest banks that dominate the Canadian banking industry. The seventh and eighth largest banks (HSBC Bank Canada and ING Bank of Canada) are, respectively, two and five times smaller than our smallest sample bank. The largest bank in our sample is the Royal Bank of Canada that ranks 32nd in the world, based on market capitalization (as of year-end 2005). With the exception of the Bank of Nova Scotia, around 90% of the banks’ revenues originate from the North-American market. Another interesting feature is that the Big Six are maintaining a very comfortable level of excess total capital over the minimum capital required by the Canadian regulator (OSFI). Moreover, the fraction of total capital required to support market risk is rather low (from 2% for Bank of Nova Scotia to 8% for Royal Bank of Canada). As a comparison, the fraction of total capital

### Table 1: Descriptive statistics

<table>
<thead>
<tr>
<th>Panel A: sample banks</th>
<th>BMO</th>
<th>BNS</th>
<th>CIBC</th>
<th>NBC</th>
<th>RBC</th>
<th>TD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market capitalization</td>
<td>28,900</td>
<td>42,600</td>
<td>24,100</td>
<td>9800</td>
<td>53,900</td>
<td>39,600</td>
</tr>
<tr>
<td>Total assets</td>
<td>297,532</td>
<td>314,025</td>
<td>280,370</td>
<td>107,598</td>
<td>469,464</td>
<td>365,210</td>
</tr>
<tr>
<td>Derivatives notional</td>
<td>1,957,456</td>
<td>775,977</td>
<td>979,013</td>
<td>373,503</td>
<td>2,816,068</td>
<td>2,265,000</td>
</tr>
<tr>
<td>Stock return</td>
<td>15.22%</td>
<td>11.45%</td>
<td>16.22%</td>
<td>16.97%</td>
<td>12.72%</td>
<td>16.47%</td>
</tr>
<tr>
<td>Stock volatility</td>
<td>19.78%</td>
<td>29.25%</td>
<td>21.16%</td>
<td>20.58%</td>
<td>17.38%</td>
<td>25.71%</td>
</tr>
<tr>
<td>Revenue from Canada</td>
<td>70.8%</td>
<td>62.7%</td>
<td>77.0%</td>
<td>95.7%</td>
<td>65.1%</td>
<td>70.7%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: regulatory capital</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tier 1 capital</td>
</tr>
<tr>
<td>Tier 2 capital</td>
</tr>
<tr>
<td>Total regulatory capital</td>
</tr>
<tr>
<td>Risk-weighted assets</td>
</tr>
<tr>
<td>Of which, market risk</td>
</tr>
<tr>
<td>Total capital ratio</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C: VaR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal model</td>
</tr>
<tr>
<td>Moving window</td>
</tr>
<tr>
<td>Conf. Level – Horizon</td>
</tr>
<tr>
<td>Start date</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of observations</th>
<th>1268</th>
<th>1012</th>
<th>1561</th>
<th>990</th>
<th>1508</th>
<th>1015</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interest rate VaR</td>
<td>11.8</td>
<td>4.6</td>
<td>3.4</td>
<td>3.5</td>
<td>14.0</td>
<td></td>
</tr>
<tr>
<td>Foreign exchange VaR</td>
<td>0.4</td>
<td>1.0</td>
<td>0.1</td>
<td>0.9</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Commodity VaR</td>
<td>3.2</td>
<td>1.7</td>
<td>1.1</td>
<td>0.6</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Equity VaR</td>
<td>3.8</td>
<td>4.3</td>
<td>5.1</td>
<td>5.1</td>
<td>7.0</td>
<td></td>
</tr>
<tr>
<td>Credit spread VaR</td>
<td>4.1</td>
<td>2.6</td>
<td>2.6</td>
<td>2.6</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Diversification effect</td>
<td>(5.5)</td>
<td>(5.0)</td>
<td>(6.0)</td>
<td>(8.0)</td>
<td>(8.0)</td>
<td></td>
</tr>
<tr>
<td>Total VaR</td>
<td>17.8</td>
<td>6.6</td>
<td>6.3</td>
<td>5.1</td>
<td>15.0</td>
<td>11.0</td>
</tr>
</tbody>
</table>

Notes: This table presents some descriptive statistics for Bank of Montreal (BMO), Bank of Nova Scotia (BNS), Canadian Imperial Bank of Commerce (CIBC), National Bank of Canada (NBC), Royal Bank of Canada (RBC), and Toronto-Dominion Bank (TD). Panel A presents for each bank its market capitalization, total assets, total notional amount of trading derivatives (interest rate contracts, foreign exchange contracts, and other), annualized stock return, annualized standard deviation of the stock return (both computed over 1999-2005), and revenues by country. Panel B presents the Tier 1 core capital, Tier 2 supplementary capital, Total regulatory capital (Tier 1 plus Tier 2 minus adjustments), the risk-weighted assets (on- and off-balance sheet credit risk plus market risk, which is separately presented), and the total capital ratio (total regulatory capital divided by risk-weighted assets). Panel C presents a description of each bank’s internal VaR model, the length of the estimation window used to extract daily VaRs, the start date and end date of our sample, the number of observations, and the VaR decomposition by risk category (bolded values).

*VaR data for NBC could not be extracted between November 1, 2002 and October 31, 2003. Unless otherwise stated, all figures are in million of Canadian dollars and as of October 31, 2005. Sources: PricewaterhouseCoopers (2006), Banks Annual Reports, and CFMRC TSE Database.*
required to support market risk in 2005 is 4.6% for Citigroup, 7.9% for HSBC, and 6.8% for UBS.

Panel C of Table 1 contains a short description of the internal VaR models used by banks. It appears that Canadian banks favor flexible, non-parametric VaR methods usually based on historical simulations. All sample banks compute and disclose one-day ahead VaRs with a 99% confidence level on a daily basis. Depending on the bank, we get daily historical data encompassing between 4 and 6 years with a total number of trading days ranging from 990 to 1561. Relying on such a long sample period is a key advantage when running a backtesting procedure since it improves the power of the statistical tests.

We plot in Fig. 1 the daily, one-day ahead, 99% VaRs and daily P&Ls for the six banks. In each graph, the grey line depicts the VaR and the black line depicts the P&L. Eyeballing Fig. 1 reveals several key features of the actual VaR and P&L data. Firstly, daily P&Ls are very volatile. Secondly, our sample offers some strong differences in P&L behavior across banks. For instance, Bank of Montreal hardly exhibits any hypothetical trading losses over our sample period and these losses are small in magnitude. Conversely, National Bank of Canada faces many hypothetical trading losses, among which some are severe. Thirdly, VaR forecasts fluctuate largely from one day to another in order to reflect daily rebalancing of the banks’

![BMO graph](image1)

![BNS graph](image2)

![CIBC graph](image3)

![NBC graph](image4)

![RBC graph](image5)

![TD graph](image6)

Fig. 1. Banks’ VaRs and profit-and-loss. Notes: This figure displays daily one-day ahead 99% VaR and profit-and-loss data as reported by six commercial banks between November 1, 1999 and October 1, 2005. The grey line depicts the VaRs and the black line depicts the profit-and-loss. All values are in million of Canadian dollars. The circles represent days on which the trading loss exceeds the VaR.
trading portfolios and volatility shocks. Fourthly, and most importantly for this study, the number of exceptions is extremely low. In our sample, two banks (Bank of Nova Scotia and National Bank of Canada) experience one exception each while the other four banks do not have any day with a loss exceeding the VaR. When compared to the expected number of exceptions in a sample of 7354 trading days of 99% VaRs (74 exceptions), the total number of exceptions (two) is surprisingly low. Our results contrast with those of Berkowitz and O’Brien (2002) in their study of the VaRs of US commercial banks. Using a 26-month sample period (vs. between 4 and 6 years in this study), they report some exceptions for all but one of their sample banks. Moreover, in their dataset, one of the banks experiences six exceptions. Although a formal statistical analysis of the actual VaRs and P&Ls is called for to answer the title question, our preliminary results suggest that the VaRs of Canadian banks are surprisingly large. We present in Table 2, some descriptive statistics on actual VaR and P&L data. The main stylized facts about VaR are its high degree of persistence and moderate skewness and kurtosis. Differently, P&Ls achieve low autocorrelation, are right-skewed, and exhibit fatter tails than the normal distribution. As already mentioned when discussing Fig. 2, there are some sharp differences between the percentage of days with loss: from a low 4% for Royal Bank of Canada to almost 35% for National Bank of Canada. Another piece of evidence pointing toward VaR overstatement is the fact that, for all banks, the average 99% VaR is much larger than the (absolute value of the) 99% percentile of the P&L distribution.

### 2.2. Backtesting

We apply standard coverage tests to the banks’ VaR to assess the accuracy of the internal VaR models. The unconditional version of the coverage test focuses on whether the actual number of VaR exceptions is equal to the expected number of exceptions (Kupiec, 1995). In a sample of 99% VaRs observed during $T$ trading days, the expected number of exceptions is $(1 - 0.99) \times T$. The null hypothesis can formally be tested using the following log-likelihood ratio test (Jorion, 2006, p. 147):

$$LR_{UC} = -2 \ln \left\{ \left(1 - p\right)^{T-X} p^{X} \right\} + 2 \ln \left\{ \left(1 - \frac{X}{T}\right)^{T-X} \left(\frac{X}{T}\right)^{X} \right\},$$

(1)

where $p$ is the target violation rate (e.g., $p = 0.01$ for a 99% VaR), $X$ is the number of exceptions, $T$ is the total number of observations, and the $LR_{UC}$ statistic is asymptotically distributed $\chi^2$ with one degree of freedom. An asymptotically equivalent unconditional coverage test is (Jorion, 2006, p. 144):

$$z_{UC} = (X - pT) / \sqrt{p(1 - p)T},$$

(2)
where the $z_{UC}$ statistic follows a standard normal distribution. One potential advantage of the $z_{UC}$ test is that it is well defined in the case that no VaR violation occurs (Campbell, 2005).

It is well known that unconditional coverage tests perform poorly in small samples. In particular, Kupiec (1995) shows that the LR UC test is prone to significant type I and type II error when applied, as suggested in Basel Committee on Banking Supervision (1996b), to 250 historical values (Jorion, 2006, p. 145). Two standard solutions are known to deal with this issue: (1) using a smaller confidence level (e.g., 95% instead of 99%); and (2) applying the coverage test to a longer historical sample period. In this paper, since the confidence level is given – all banks use a 99% confidence level – we consider a relatively long sample period spanning between 4 and 6 years. Another well known limitation of the unconditional coverage test is that it ignores exception clustering. Christoffersen (1998) develops a conditional coverage test, LR CC, that formally accounts for clusters of exceptions:

$$
\text{LRCC} = \text{LR}_{UC} + \text{LR}_{IND} 
$$

$$
\text{LR}_{IND} = -2 \ln \left\{ (1 - \frac{1}{T})^{T_{00} + T_{01}} \left( \frac{X}{T} \right)^{T_{00} + T_{11}} \right\} 
+ 2 \ln \left\{ (1 - \pi_0)^{T_{00}} \pi_0^{T_{01}} (1 - \pi_1)^{T_{10}} \pi_1^{T_{11}} \right\},
$$

where $T_{ij}$ is the number of days in which state $j$ occurred in one day while it was at $i$ the previous day and $\pi_i$ is the probability of observing an exception conditional on state $i$ the previous day, $\pi_0 = T_{00}/(T_{00} + T_{01})$ and $\pi_1 = T_{11}/(T_{10} + T_{11})$, and the LR CC statistic is asymptotically distributed.

Fig. 2. Measure of risk overstatement. Notes: This figure presents the exception rate (left axis) and the number of exceptions (right axis) obtained by applying a discount factor to the actual disclosed banks’ VaR (DVaR). The discounted VaR is defined as $\text{DVaR} \times (1 - \rho)$, with $\rho \in [0, 1]$. The bank-specific risk-overstatement coefficient $\hat{\rho}$ is defined as the minimum $\rho$ such that the adjusted time series $\text{DVaR} \times (1 - \rho)$ is not rejected by the unconditional coverage test at the 10% confidence level. The grey shaded area corresponds to the 10% confidence level non-rejection range. The non-rejection range is [8–18] for BMO, [6–15] for BNS, [10–22] for CIBC, [6–15] for NBC, [10–21] for RBC, and [6–15] for TD.
Notes: This table presents for each bank the unconditional coverage tests, LR UC and $z_{UC}$, presented in Eqs. (1) and (2) and the conditional coverage test, LR CC, presented in Eq. (3). The tests are applied to the bank’s VaR forecasts, as well as to two other VaR samples based on historical simulation and a GARCH(1,1) model, respectively. For the two benchmark models, we estimate the VaR at time $t$ using a moving window of 250 days covering the period from $t - 250$ to $t - 1$. An implication of the one-year estimation window is that the sample period used to compute the test statistics for the benchmark models is one year shorter than the one used to compute the test statistics for the banks’ VaR. LR UC and LR CC are asymptotically $\chi^2$ with one and two degrees of freedom, respectively, and $z_{UC}$ is asymptotically standard normal. Bold figures denote occurrences when the null hypothesis is rejected at the 95% confidence level. The $\chi^2$ critical value is 3.84 with one degree of freedom and 5.99 with two degrees of freedom and the standard normal two-tailed critical value is 1.960. The sample period is November 1, 1999 and October 31, 2005 (or part of it for some banks, see Table 1).

<table>
<thead>
<tr>
<th>Bank</th>
<th>Banks’ VaR</th>
<th>Historical simulation</th>
<th>GARCH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Actual exceptions</td>
<td>LR UC</td>
<td>$z_{UC}$</td>
</tr>
<tr>
<td>BMO</td>
<td>0</td>
<td>–3.579</td>
<td>11</td>
</tr>
<tr>
<td>CIBC</td>
<td>0</td>
<td>–3.971</td>
<td>8</td>
</tr>
<tr>
<td>NBC</td>
<td>1</td>
<td>13.296</td>
<td>–2.843</td>
</tr>
<tr>
<td>RBC</td>
<td>0</td>
<td>–3.903</td>
<td>11</td>
</tr>
<tr>
<td>TD</td>
<td>0</td>
<td>–3.202</td>
<td>10</td>
</tr>
</tbody>
</table>

$\chi^2$ with two degrees of freedom.\textsuperscript{13} Specifically, the null hypothesis in the independence test $\text{LR}_{\text{IND}}$ states that the likelihood of a violation on a given day does not depend on whether or not a violation occurred on the previous day. Unlike the unconditional coverage tests that can be implemented with aggregate data, the latter test requires a complete time series of daily VaR and P&L data.

We present the value of the $\text{LR}_{\text{UC}}$, $z_{\text{UC}}$, and $\text{LR}_{\text{CC}}$ statistics in Table 3. All the tests provide clear evidence that the banks’ VaRs are excessively large (see first column). The hypothesis that states that the actual number of exceptions is equal to the expected number of exceptions is systematically rejected at the 95% confidence level. This result was expected for the four banks that did not experience any exception but it also holds for Bank of Nova Scotia and National Bank of Canada, which both experienced one exception.

To further assess the accuracy of the proprietary models, we compare banks’ VaRs to VaR forecasts given by two benchmark models. The two approaches that we follow are based on historical P&L data only. As a result, our information set is smaller than the one of the bank since we do not know the bank’s trading positions, which a priori would bias against our benchmark models outperforming the internal model. We estimate the two benchmark models using a one-year historical dataset of P&L data that we update daily.

The first approach is the historical simulation method (Linsmeier and Pearson, 2000). In a recent survey, Péronon and Smith (2007) show that historical simulation is at the heart of most proprietary models currently used by commercial banks. For each sample day, we sort in ascending order the P&Ls over the past 250 trading days to arrive at an observed distribution of P&L values. Since there are 250 observations, this means that 2.5 ≈ 3 losses will be larger than the VaR figure.\textsuperscript{14} If we denote the P&L on day $t$ by $R_t$, the historical simulation-based VaR forecast is:

$$\text{VaR}^\text{HS}_t = -\min_{|\Delta| \leq 250} R_t,$$

where the $\min_{|\Delta|}$ operator denotes the third smallest value from a set of values. The second approach is based on the GARCH model, which (unlike historical simulation) has the desirable feature of removing clusters in the VaR exceptions. We rely on the simplest possible GARCH specification, which is a GARCH(1,1) model with a basic mean equation:\textsuperscript{15}

$$R_t = \alpha + \epsilon_t,$$

$$\sigma_t^2 = \beta_0 + \beta_1 \sigma_{t-1}^2 + \beta_2 \epsilon_{t-1}^2,$$

$$\text{VaR}^{\text{GARCH}}_t = -(\bar{R}_t - 2.33 \times \bar{\sigma}_t),$$

where $\bar{R}_t$ is given by Eq. (6) and $\bar{\sigma}_t$ by Eq. (7).

\textsuperscript{13} Following Christoffersen (2003, p. 188), when implementing the $\text{LR}_{\text{IND}}$ test with $T_{11}=0$, we simplify the second term in Eq. (4) as $2 \ln \left(1 - \pi T_{10} / \pi_0 \pi T_{10}^{1/2} \pi T_{11}^{1/2} \pi T_{11}^{1/2} \right)$. Furthermore, if $T_{10} + T_{11} = 0$, $\pi_1$, $\text{LR}_{\text{IND}}$, and in turn $\text{LR}_{\text{CC}}$ are not defined. However, in our empirical study, this arises only with banks’ VaR figures but not with historical simulation or GARCH-based VaRs. The alternative conditional coverage test developed by Christoffersen and Pelletier (2004), which is based on the time between two exceptions, is also subject to the same problem. As noted by a referee, the finding that there are no exceptions seems to indicate that the independence assumption might be violated. Although there is a positive probability to end up with no exceptions at all after one year with violations independently distributed through time, this probability is likely to be much larger with dependence.

\textsuperscript{14} Our version of the historical simulation method is a simplified one. Standard historical simulation requires, for each particular position (unknown to us), the identification of all market risk factors. The method yields a hypothetical distribution of the changes in each position value, and after aggregation, in the trading portfolio value, from which the VaR can be computed.

\textsuperscript{15} GARCH modeling is more general than the standard Riskmetrics variance model (J.P. Morgan, 1996; Christoffersen, 2003, p. 22). Following Berkowitz and O’Brien (2002), we also estimate an ARMA(1,1)–GARCH(1,1) model, as well as an AR(1)–GARCH(1,1). We obtain similar results and therefore do not present the results here.
A key difference between historical simulation-based VaRs and GARCH-based VaRs is that the former exhibit a relatively sluggish dynamic and the latter are very reactive to actual P&L shocks. The two benchmark models yield smaller (average and median) VaRs than banks’ internal models, which is consistent with US evidence in Berkowitz and O’Brien (2002). A direct consequence of this scaling effect is that the number of exceptions generated by each benchmark model is significantly larger than the actual number of exceptions. An important result in Table 3 is that neither the unconditional coverage tests nor the conditional coverage test ever reject the two benchmark models. This contrasts with banks’ internal models that are systematically rejected by the coverage tests. We interpret this set of results as additional evidence about the banks’ excess of conservatism in their VaR estimates.

2.3. Measuring VaR overstatement

We aim in this subsection to quantify the VaR overstatement of each bank. To do so, we formally define the one-day ahead VaR with a confidence level \( c = 1 - \rho \) as the quantity such that:

\[
\text{Prob}(R_{t+1} < -\text{VaR}_t | I_t) = \rho,
\]

(9)

where \( I_t \) is the information set at time \( t \). In case of VaR overstatement, the disclosed VaR (DVaR) is such that \( \text{DVaR}_t > \text{VaR}_t = \text{DVaR}_t \times (1 - \rho) \), where \( \rho \) is a risk-overstatement (RO) coefficient, \( \rho \in [0, 1] \), and:

\[
\text{Prob}(R_{t+1} < -\text{DVaR}_t | I_t) < \rho,
\]

(10)

or, equivalently,

\[
\text{Prob}(R_{t+1} < -\text{DVaR}_t \times (1 - \rho) | I_t) = \rho.
\]

(11)

As an illustration, let us consider a given bank that knows that it has a 1% change to lose at least $50 million with a one-day horizon, i.e. its one-day ahead 99% VaR is $50 million. If the bank discloses a VaR equal to $60 million (DVaR = 60), it implies that \( \rho \) is 0.167 or, in other words, that the bank intentionally inflates its VaR by 20% or \( 0.167/(1 - 0.167) \). Like for all our sample banks, the VaR model of a risk-overstating bank is going to be rejected by the coverage test since the model does not yield enough exceptions. However, if the disclosed VaRs were sufficiently reduced, the number of exceptions would increase and the null hypothesis \( \rho = 0.01 \) would not be rejected anymore. Using a sample of historical disclosed VaRs and P&Ls, we can estimate the value of the risk-overstatement coefficient. It corresponds to the minimum \( \rho \) such that the adjusted time series \( \text{DVaR}_t \times (1 - \rho) \) is not rejected by the unconditional coverage test at a given confidence level.

For each sample bank, we compute the LRUC statistics for \( \text{DVaR}_t \times (1 - \rho) \) for \( \rho \in [0, 1] \) with a 0.001 increment and we plot its value in Fig. 2. Taking Bank of Nova Scotia as an example, we see that \( \text{DVaR}_t \times (1 - \rho) \) is not rejected at the 10% confidence level by the unconditional coverage test for \( \rho \in [0.526, 0.670] \) since the number of exceptions remains within the non-rejection range [6,15]. The risk overstatement estimates go from a moderate 19% for National Bank of Canada to an extremely high 79% for Bank of Montreal and Royal Bank of Canada. Our empirical results point toward a strong and systematic VaR overstatement, which is consistent with the US evidence reported in Berkowitz and O’Brien (2002). However, the degree of conservativeness is clearly larger for Canadian banks. Our primary findings show that the internal models approach does not meet its main objective, which is to “lead to capital charges that more accurately reflect individual banks’ true risk exposure” (Hendricks and Hirtle, 1997, p. 3).

2.4. Cost of overstating VaR

While over-reporting VaR can appear attractive at first sight (especially to regulators), it is costly for the bank and, potentially, for the economy. In this subsection, we first quantify the cost in terms of excess regulatory capital induced from an inflated VaR. Under the framework defined in Basel Committee on Banking Supervision (1996a), a bank’s Market Risk Charge (MRC) directly depends on its 99% VaR.

\[
\text{MRC}_t = \max \left( 3\sqrt{10} \frac{1}{60} \sum_{i=1}^{60} \text{VaR}_{t-i}(99\%, \text{1 day}), \right.
\]

\[
\sqrt{10}\text{VaR}_{t-1}(99\%, \text{1 day}) \right),
\]

(12)

where 3 is an ad-hoc multiplicative factor imposed by regulators. In case of risk overstatement, \( \text{VaR}_t \) is replaced by \( \text{DVaR}_t = \text{VaR}_t/(1 - \rho) \) and the market risk charge becomes:

\[
\text{MRC}^{\text{RO}}_t = \max \left( 3\sqrt{10} \frac{1}{60} \sum_{i=1}^{60} \frac{\text{VaR}_{t-i}(99\%, \text{1 day})}{1 - \rho}, \right.
\]

\[
\sqrt{10} \frac{\text{VaR}_{t-1}(99\%, \text{1 day})}{1 - \rho} \right),
\]

(13)

\[
\text{MRC}^{\text{RO}}_t = \frac{1}{1 - \rho} \text{MRC}_t.
\]

(14)

\footnote{For simplicity, we neglect the second component of the market risk charge, which is the specific risk charge that aims to account for the idiosyncratic risk of the trading portfolio. We replace the 10-day ahead VaR by the 1-day ahead VaR using the square-root of time rule.}
We quantify the annual dollar cost (DC) of reporting inflated VaRs – and in turn, of having too large a regulatory capital – by defining:

$$DC = (\text{MRC}^{\text{RO}} - \text{MRC}) \times \text{ROA},$$

where ROA is the bank’s return on assets which is used as a proxy for the opportunity cost of the regulatory capital.

We compute the dollar cost for each bank using the value of $\rho$ estimated in Fig. 2, the market risk charge ($\text{MRC}^{\text{RO}}$) extracted from Table 1, and the average ROA between 1999 and 2005. We exhibit in Fig. 3 the estimated values of the DC variable. Although the excess market risk charges, $\text{MRC}^{\text{RO}} - \text{MRC}$, are large numbers (from CAD58.6 million for National Bank of Canada to CAD1,230.6 millions for Royal Bank of Canada), the magnitude of the dollar cost of over-reporting VaR is rather small (from CAD0.4 million for National Bank of Canada to CAD8.3 millions for Royal Bank of Canada).

Besides the cost directly borne by the banks, there are some other potentially negative implications induced by VaR over-reporting. Firstly, an important asset-pricing implication of bank’s exaggeration of their own level of risk is that banks appear more risky than they actually are. Consequently, if investors truly believe that disclosed VaRs are accurate measures of individual banks’ risk exposure, the expected return required by investors will be higher than necessary, which in turn, distort current stock prices. Moreover, VaR overstatement may also lead to inefficient portfolio allocations. An investor incorporating VaR into his estimates of banks’ standard deviation would tend to systematically underinvest in VaR-overstating banks. Secondly, the inefficient allocation of the resources within the bank, i.e. excessive regulatory capital, may prevent some attractive projects being funded, which has a detrimental effect on the growth of the economy.

### 2.5. Causes of inflated VaRs

In the previous sections, we have shown that, on average, VaRs disclosed by commercial banks are higher than they need to be. A first potential explanation is that all commercial banks’ internal models are unable to precisely measure market risk. This is however unlikely to be the case since, as shown in Section 2.2, it is relatively easy to develop an acceptable VaR model. Furthermore, it seems unlikely that banks’ risk managers would systematically make mistakes in the same direction, i.e. overestimation.

A potential source of VaR overstatement is the VaR aggregation process. Indeed, commercial banks compute VaR estimates separately for each business line (see Bierkowitz et al., 2006) and/or for each risk category (see Bierkowitz and O’Brien, 2002; and our Table 1). Since banks are required to disclose a comprehensive, firm-level VaR measure, they have to combine sub-group VaRs to end up with...
a single aggregative measure. Berkowitz and O’Brien (2002, p. 1109) report that, over the period 1998–2000, US commercial banks obtained their aggregated VaR by simply summing sub-group VaR.\(^{21}\) The practice of neglecting correlation reflects the widely held concern by industry practitioners that “correlations go to one” in a crisis but it has the immediate effect of inflating aggregate VaR estimates. For instance, we see in Table 1 – Panel C that the correlation effect for National Bank of Canada is comparable in size with the final VaR. As a result, ignoring diversification would lead this firm to double its VaR. Another example of the importance of the diversification effect in aggregating VaR can be found in Berkowitz et al. (2006) who show that, for an international bank, exceptions do not occur at the same time in different business lines. Although all sample banks systematically account for correlation effects when calculating their market risk (see Table 1), they tend to underestimate the diversification effect.\(^{22}\) For instance, Bank of Montreal asserts in its annual report for 2003 (p. 49): “In determining VaR, we take only partial account of the correlation and offsets that may exist between certain portfolios and classes of risk [...]. A more conservative measure of market risk is therefore calculated than would otherwise be the case”.

Another reason for exaggerating VaR figures is that banks do not want to put their reputation at risk. As a result, they are deliberately extremely cautious when reporting their VaR. Indeed, the market is likely to severely punish any bank unable to correctly assess its exposure to market risk. This form of the “prudent man rule” is likely to explain a substantial part of the risk overstatement brought to light in this paper. This argument has a good chance of applying to the Big Six since they are all well-capitalized commercial banks that can easily afford to overstate their VaR and, in turn, their regulatory capital.

Finally, risk officers may also have an incentive to voluntarily increase the VaR confidence level just so they do not attract attention (internally and externally). This allegation has been validated through discussions with risk officers, regulators, and consultants and is also put forward by Jorion (2006, p. 155).

3. Conclusion

Based on the overwhelming evidence presented in this paper, the answer to the title question is clearly positive. Over our sample period covering 1999–2005, commercial banks have displayed a tendency toward conservatism when setting their VaR. For some sample banks, this bias is found to be extremely severe. However, large commercial banks have not been overstating their VaR over the entire post 1996 Basel Accord amendment period. For instance, Jorion (2006, p. 144) mentions that J.P. Morgan experienced 20 exceptions in 1998, which is far more than the 13 or so it expected with a 95% confidence level. In the past, some VaR understatements were due to inaccurate risk assessments. Currently, modern banks overstate their VaR but the reason is not that they are unable to correctly measure their market risk. We claim that market risk overstatement comes from the fact that banks want to be very cautious when setting their VaR, and that they underestimate the effect of diversification when aggregating VaR across business lines and/or risk categories. The conclusion of our study indicates that modern banks are moving forward on the learning curve of VaR determination.

While the present study analyzes actual VaR of large and well-capitalized banks, future research could apply our methodology to VaR disclosed by smaller banks, as well as VaR of banks operating in different regulatory environments. Another useful complement to the present study would be to compare levels of VaR overstatement across aggregation levels, e.g. trader/trading desk/business line/bank. Furthermore, our methodology could also be extended to allow for time variation in the degree of VaR overstatement.

Acknowledgements

We thank Jeremy Berkowitz, George Blazenko, Dennis Chung, Ramo Gencay, Rob Grauer, Thomas Gilbert, Campbell Harvey, Robbie Jones, Peter Klein, Pascal Lavergne, Fred Shen, and two anonymous referees for their comments and suggestions. We are particularly grateful to Peter Christoffersen and Daniel Smith for their very helpful comments and to Phil Goddard for his help with the data extraction. We thank several officers from the Office of the Superintendent of Financial Institutions Canada and risk officers at major Canadian banks for their feedback. Pérongon thanks the Social Sciences and Humanities Research Council of Canada for financial support.

Appendix A. Step-by-step user’s guide for the MATLAB-based data extraction method

1. Convert the original graph from the annual report (available as a PDF file) into a JPG file.
2. Import the JPG file into MATLAB and define it as an image called, for instance, im. Command: im = imread(‘name.jpg’).

---

\(^{20}\) Aggregating sub-group VaR is problematic for different reasons. First, VaR is not a subadditive risk measure (see Artzner et al., 1999), i.e. VaR(A + B) is not always smaller than VaR(A) + VaR(B), which neglects the fact that risks usually diversify away when put together. Second, the number of risk factors affecting trading revenues can be very large. For example, in its 2004 Financials & Form 10-K, Citigroup states that its “Value-at-Risk is based on the volatilities of, and correlations between, approximately 250,000 market risk factors”.

\(^{21}\) Nowadays, most US banks account for diversification effects.

\(^{22}\) An alternative interpretation is that the diversification adjustment is accurate but the sub-group VaR components are inflated. However, this latter assertion would be in opposition with empirical evidence on business-line VaR reported by Berkowitz et al. (2006) for an unidentified international bank. Indeed, they show that disaggregated VaR are relatively accurate.
3. Display the image in MATLAB. Command: `image(im)`.
4. Convert the graph scale into a MATLAB scale. For instance, the zero value on the vertical axis of the graph corresponds to a value of 100 in MATLAB, 10 corresponds to 80, 20 corresponds to 60, etc. This implies a conversion scale factor of $s = 2 = (100 - 80)/(10 - 0) = (80 - 60)/(20 - 10)$.
5. Add vertical lines on the image (Edit, Axes Properties, Show Grid, and manually enter the coordinates where the lines have to be in ‘Ticks’). The horizontal distance $d$ between two adjacent lines is defined using two successive clear-cut data points. Find the coordinate value on the horizontal axis for the first and last data point, which we call $f$ and $l$. The number of vertical lines is $(l - f)/d$ and the number of observations is $1 + (l - f)/d$. Note that the fact that the exact number of data points is not required to be known is a key advantage of our data extraction algorithm.
6. Zoom in and click on each data point. By doing so, we capture the two-dimensional coordinates of each data point. Command: `data = ginput(n)`, where $n$ is the number of data points to be extracted. Then, sequentially click on the data points. The 2-D coordinates are automatically stored into the $(n \times 2)$ data matrix.
7. Convert the MATLAB vertical coordinates (second column of the `data` matrix) into graph coordinates. For each data point, compute $(zero\ coordinate\ value - point\ coordinate\ value)/s$. As a refinement, we can plot the extracted data in EXCEL and superimpose the graph of the extracted data with the original graph. This can be done by right clicking on the graph of the extracted data and Format Chart Area/Area: none and Format Plot Area/Area: none. If necessary, we can manually adjust the extracted series until reaching a perfect match between the two lines.

**References**