Efficiency, Economies of Scale and Scope of Large Canadian Banks*

FIRST DRAFT APRIL 2004

Jason Allen
Department of Economics
Queen's University

and

Ying Liu
Department of Monetary and Financial Analysis
Bank of Canada

June 1, 2004

ABSTRACT

This paper measures economies of scale and scope of Canada's six largest banks and their cost efficiency over time. Using a unique panel of data from 1983q1 to 2003q3, we estimate pooled translog cost functions and derive measures for relative efficiency, and economies of scale and scope. The disaggregation of the data allows us, for the first time in the literature on Canadian banks, to model banks as producing multiple outputs, including non-traditional activities. Given the long sample period of our dataset, we are able to incorporate possible technological and regulatory changes in the banks' cost functions, as

*The authors thank Allan Crawford, Stefan Dodds, Geoffrey Dunbar, Walter Engert, Kim Huynh, Ian Keay, Jeremy Lise, Darcey McVanel, and Greg Tkacz, as well as participants at a Bank of Canada workshop. All errors are our own. The views expressed in this paper are those of the authors. No responsibility for them should be attributed to the Bank of Canada. Jason Allen thanks the Bank of Canada for their hospitality and facilities while writing this paper.

well as time-varying bank-specific effects. Our model leads us to reject constant returns to scale. We also find mixed evidence regarding economies of scope. These findings seem to suggest there are potential scale benefits in the Canadian banking industry. We also find that technological and regulatory changes have had significant positive effects on the banks' cost structure.

Keywords: banking, economies of scale and scope, panel data
JEL classification: G21, D24, C33

I. Introduction

This paper measures the economies of scale and scope, and efficiency in the Canadian banking industry. These measures provide important insights not only to managers in their process of making operational decisions, but also to policy makers in their debate on regulatory issues such as mergers. Using a unique panel of data of the six largest Canadian banks from 1983q1 to 2003q3, we estimate pooled translog cost functions and derive measures for efficiency, and economies of scale and scope. Economies of scale allow us to assess whether "larger is better", based on existing technology. Economies of scope identify potential cost savings/dissavings in the combination of certain product lines.

The disaggregation of the data allows us, for the first time in the literature on Canadian banks, to model banks as producing multiple outputs. Included as bank output are non-interest related activities such as deposit account services, security underwriting and wealth management. Non-interest activities have rarely been studied in the literature. We proxy these activities by an asset-equivalent-measure of non-interest income Boyd and Gertler (1994).

The analytical framework used in this study is the flexible translog cost function, assuming the intermediation approach in bank production (c.f. Murray and White (1983)). Three econometric models are estimated: a time-varying fixed-effect panel model, a stochastic cost efficiency frontier model estimated by maximum likelihood (MLE), and a system of unrelated regressions (SUR) using GLS. We also include a fixed effect in the SUR model to represent inefficiency.[1] Measures of economies of scale and scope are calculated from the derivatives of cost with respect to output. Under the fixed effect and stochastic frontier model, we are also able to generate measures of relative cost efficiency among banks.[2]

[1] We initially estimated a model with a composite error term, i.e. no fixed effect using the SUR approach. This was the typical method for estimating cost functions and the coefficient estimates are efficient if variables are exogenous. For example see Murray and White (1983). However, using the Hausman test we reject this assumption and favour a fixed effect approach.

[2] Two components of efficiency can be distinguished: technical efficiency, the ability to obtain maximum output from a given set of inputs, and allocative efficiency, the skill to use the inputs in optimal proportions, given their respective prices and the production technology. The two can be combined to provide a measure of economic efficiency, or, when cost instead of production is considered, cost efficiency.

Given the long sample of our dataset, we are able to identify possible technological changes over time. Here, technological change is a broad term which includes financial innovations, changes in the competitive nature of banks, and demographically led changes in household portfolios. Indeed, Freedman and Goodlet (1998) note that banks have been undergoing significant technological changes in recent years which affect the way services are provided, the instruments used to provide services and the nature of the financial service providers. These changes include the adoption of electronic processing of transactions, the development of new instruments and products, the spread of ATMs and internet banking and better risk management techniques. We model technological changes in two ways: imposing common technological trends on the cost function or assuming technological changes affect bank-specific progress.

Another important factor that may have affected the banks' cost structure over time is regulatory changes. Calmès and Liu (2003) argue that changes to the Bank Act in 1987, 1992 and 1997 may have encouraged the trend towards direct financing, i.e. financing done at financial markets rather than through financial intermediaries. At the same time, banks have been increasingly involved in non-traditional, typically market-oriented, activities. We attempt to model the potential impacts of such changes in the sources of a bank's income on its cost structure by introducing regulatory variables.

Our results suggest that we reject constant returns to scale. Depending on the model, our results suggest that banks can reduce average cost by five to 20 per cent by doubling their output. Similar to those in the banking literature, our findings on economies of scope are mixed. We also conclude that both technological and regulatory changes have had positive effects on the banks' cost structure.

The paper is organized as follows: Section II introduces the literature on efficiency, economies of scale and scope of financial institutions. Section III presents our model. Section IV discusses the econometric issues and the models that we estimate. Section V provides a detailed description of the data. Section VI presents our estimation results. Section VII summarizes our findings.

II. Literature Review

Economies of scale and scope in financial institutions has a long history. While studies have been done for different types of financial institutions in different countries, very few focus on Canadian institutions. In general, most studies find only small economies of scale in a firm's cost structure. In those studies that find evidence of economies of scale, the measured economies of scale seem to be stronger in small to medium sized firms than for large firms.[3] However, more recent studies find stronger evidence of economies of scale in large U.S. commercial banks in the 1990s (Berger and Mester (1999), and Stiroh (2000)). On the other hand, evidence about economies of scope is mixed.[4]

Two recent studies on Canadian financial institutions address economies of scale assuming a Cobb-Douglas cost function. Using a panel data set of 25 Canadian trust companies for the years 1985-1988, Breslaw and McIntosh (1997) show that the scale function of these trust companies is convex with respect to firm size. Assuming a Cournot oligopoly and using a time-series data set of the Big Five[5] from 1976 to 1996, McIntosh (2002) finds significant increasing returns to scale among these banks.

A limitation of the Cobb-Douglas framework is that banks are assumed to produce one (composite) output with the same inputs and technology. This assumption is debatable (and rejected in our sample) given that banks are diversified in their business lines. Bank output can no longer be proxied by simple measures such as total assets, total revenue or total number of loans.

The Cobb-Douglas framework also has an overly-restrictive functional form. Lawrence (1989) demonstrates that the non-rejection of the Cobb-Douglas technology by previous studies results from an ad hoc specification that excludes the possibility of multiple product cost complementarity. Applying this to British Columbia Credit Unions, Murray and White (1983)

[3]See Ferrier and C.A.K.Lovell (1990) for U.S. banks, Rime and Stiroh (2003) for Swiss universal banks, Rezvanian and Mehdian (2002) for Singaporian banks.

[4]Studies that find economies of scope include Murray and White (1983). Studies that reject economies of scope include Berger and D.B.Humphrey (1985).

[5]Royal Bank Financial Group, Bank of Montreal, Canadian Imperial Bank of Commerce, TD Bank Financial Group, and Bank of Nova Scotia

show that none of the restrictive production conditions commonly imposed by researchers using the Cobb-Douglas framework provides a valid representation of the technology of the firms they studied. Instead, the authors propose to use a translog specification which captures the heterogeneous nature of a bank's intermediation activity. Using data from 1976 to 1977, Murray and White find that most of the credit unions in the sample experience significant increasing returns to scale and there is some evidence of economies of scope in their mortgage and other lending activities.

The translog cost function was first proposed by Christensen and Lau (1971). Schmidt and Lovell (1979) show that under the cost minimization assumption, a firm's stochastic production frontier can be written as a cost function. Diewert and R.J.Kopp (1982) show that any frontier cost function, such as the translog function, can be derived without knowing its underlying production function. The translog specification is often used to provide a numerical efficiency value, called X-efficiency, and ranking of firms. The X-efficiency of a bank is measured as its cost level compared to that of the best-practice banks of similar size (the frontier firm), controlled for its types of banking activity and the input prices it faces. Inferences regarding scale and scope economies of banks are drawn from the derivative of a bank's cost with respect to its output and the cross elasticity between outputs, respectively. The specification is often applied to a set of cross-sectional data on banks and estimated for several years. The parameters for economies of scale and scope are averaged over the sample years. Berger and Humphrey (1997) provide a detailed survey on the literature.

While most studies using the translog cost function examine U.S. or European banks, to our knowledge, no study has applied it to Canadian banks. This is partly due to the limited number of banks in Canada and partly due to the unavailability of individual bank data to the public.[6] However our panel has a significantly long time dimension and a sufficient cross-sectional dimension to conduct a detailed study of the Canadian banking industry.

A major difficulty in dealing with large-T, small-N panels is that the assumption of constant firm effects over time is most likely violated. Cornwell and Sickles (1997) propose to solve the problem by replacing the coefficents on firm effects with a flexibly parameterized function

[6] As of July 2003, there were 17 domestic chartered banks and 51 foreign banks or subsidiaries in Canada.

of time. Applying this to eight U.S. airlines from 1970q1 to 1981q4, the authors construct an efficiency frontier for the firms and find that the firms become more efficient over time. We follow the approach of Cornwell and Sickles (1997) to allow bank specific effects to vary over time.

A further contribution of this paper is the inclusion of non-traditional activities. The aforementioned studies mostly measure bank output by their traditional activities, such as loan-generating and security investment. However, banks have been trending towards non-traditional activities such as depositor services, underwriting, and wealth management. Excluding these activities will result in a misspecified cost function and thus incorrect inference about economies of scale and scope.

Clark and Siems (2002) apply the asset equivalent measure of off-balance-sheet (OBS) activities proposed by Boyd and Gertler (1994) to measure the impact of off-balance sheet (OBS) activities on the efficiency measure of banks. The authors find that OBS activities are significant in explaining cost efficiency. The idea of the asset equivalent measure is to capitalize the bank's noninterest income to proxy the assets required to produce such revenue. Boyd and Gertler (1994) also argued that the credit equivalent measure proposed by the Basel Committee often underestimates the off-balance-sheet assets of banks.[7] We follow the framework of Clark and Siems (2002). In addition to the efficiency of banks, we investigate the economies of scale of each type of bank output as well as their compatibility.

[7]Under the current reporting requirement, OBS activities are allocated into four broad risk categories that carry conversion factors of 100 per cent, 50 per cent, 20 per cent and 0 per cent.

III. Models

A. Cost Minimization

We assume a bank i, $(i = 1,...,N)$ is a cost minimizing entity who produces output $Q = (Q_1,...,Q_m) \in R^m_+$ using inputs $X = (X_1,...,X_k) \in R^k_+$ at prices $W = (W_1,...,W_k) \in R^k_+$ subject to a production constraint, $F(Q,X)$.[8]

$$\min C = \sum_{j=1}^{k} W_j X_j(Q,W)$$

subject to a production constraint

$$F(Q,X) = 0$$

Possible environmental variables and proxies for technological change are included in $G = (G_1,...,G_L) \in R^L_+$. Consistent with most of the firm-efficiency literature we estimate a multiproduct translog cost function.[9] The function is assumed to be positive for all positive prices and outputs, homogeneous of degree one, monotonic, and concave in prices. The total cost function, $C(Q,W)$ can be defined in terms of the unit cost function $c(w)$ as follows: $c(q,w) = q + c(w)$. A second order Taylor expansion around the log of output and prices gives the following cost function. All lower case variables are in logarithms.

$$c(q,w) = \alpha_0 + \sum_{l=1}^{m} \alpha_k q_l + \sum_{j=1}^{k} \beta_j w_j + \frac{1}{2} \sum_{l=1}^{m} \sum_{j=1}^{m} \sigma_{lj} q_l q_j + \sum_{l=1}^{m} \sum_{j=1}^{k} \gamma_{lj} q_l w_j + \frac{1}{2} \sum_{l=1}^{k} \sum_{j=1}^{k} \delta_{lj} w_l w_j + \sum_{l=1}^{L} \theta_l G_l + \xi + \varepsilon \quad (1)$$

At this time we do not subscript the variables across time or cross-section. We do this explicitly because subscripts depend on the methodology employed. The assumption on the

[8] Shaffer (1993) tests for banking competition in Canada between 1965 and 1989 and cannot reject input price-taking behaviour.

[9] The Fourier functional form has been shown by Mitchell and Onvural (1996) to perform better than the translog function when there is a wide variety of banks. This is clearly not the case in our sample and thus we estimate the simpler of the two models.

distribution of ξ and ε is also dependent on the methodology. Economic theory imposes certain restrictions on the parameters: the cost function is homogeneous if: $\sum_j^k \beta_j = 1$, $\sum_j^k \gamma_{lj} = 0$, and $\sum_j^k \delta_{lj} = 0$. For simplicity we also impose $\sigma_{lj} = 0$.This allows us to re-write the problem as,[10]

$$\log(c/w_1) = \sum_{j=2}^{k} \beta_j \log(w_j/w_1) + \sum_{l=1}^{k-1} \sum_{j=l+1}^{k} \delta_{lj}[w_l w_j - 0.5 * (w_l^2 + w_j^2)] + \sum_{l=1}^{m} \sum_{j=2}^{k} \gamma_{lj} q_l \log(w_j/w_1) + \alpha_0 + \sum_{l}^{m} \alpha_l q_l + \sum_{l=1}^{L} \theta_l G_l + \xi + \varepsilon \quad (2)$$

It is important to note at this point that we gain efficiency by including input cost share equations in the model. The cost share of input j is given by S_j and is derived using Shephard's lemma. The cost share equations add information without adding parameters to the multivariate regression. Since shares must sum to unity, $j-1$ input share equations are specified for a system of j shares.

$$S_j = \frac{\partial \log(C)}{\partial \log(w_j)} = \beta_j + \sum_l \delta_{lj} w_l + \sum_l \gamma_{lj} q_l, \qquad j = 1, ..., m \quad (3)$$

Imposing the appropriate parameter restrictions gives us[11]:

$$S_j = \beta_j + \sum_{l=2}^{m} \delta_{lj} \log(W_l/W_j) - \sum_{l=1}^{m} (\gamma_{l2} + \gamma_{l3}) q_l \quad (4)$$

The appropriate cross-equation restrictions are imposed. Given parameter estimates we also derive own- and cross-price elasticities (ν_{lj}) and Allen partial elasticities of substitution (Ω_{lj}).

$$\nu_{lj} = \frac{\gamma_{lj}}{S_l} + S_j - g_{lj} \quad \textit{and} \quad \Lambda_{lj} = \frac{\gamma_{lj}}{S_l S_j} + 1 - \frac{g_{lj}}{S_j}$$

where $g = 1$ for $l = j$, and zero otherwise. Concavity of the cost function is checked by assuring that the eigenvalues of the matrix Λ_{lj} are non-positive at every data point.

[10] Dividing the cost function by W_1 is more practical for estimation purposes. The coefficients have the same interpretation.

[11] Note that symmetry imposes $\gamma_{lj} = \gamma_{jl}$.

B. Economies of Scale and Scope and Technological Changes

Our main interest lies in the economies of scale and scope of the Canadian banking industry. Economies of scale are defined as:

$$\zeta = \Big(\sum_{l=1}^{m} \alpha_l + \sum_{l=1}^{m} \sum_{i=2}^{m} \gamma_{ij} \log(W_l/W_j) \Big)^{-1}$$

and economies of scope is approximated as[12]:

$$\eta_{lj} = -\alpha_l \alpha_j$$

There are increasing returns to scale if $\zeta > 1$, constant returns to scale if $\zeta = 1$, and decreasing returns to scale if $\zeta < 1$. There are economies of scope if $\eta < 0$ and diseconomies of scope of $\eta > 0$. Economies of scale inform us of the cost savings/dissavings when a bank increases its output. Economies of scope inform us of the cost savings/dissavings from combining product lines.

Given the long dimension of our sample, we also consider capturing possible technological changes over time. We achieve this in two ways. First, we assume technological change affects the cost function directly, i.e. banks are subject to the same technological shocks over time. We proxy such shocks by including a time trend and time trend in G_l. The rate of technological change is given by

$$T^* = -\frac{\partial C}{\partial t} = -(\theta_1 + 2\theta_2)$$

and change is progressive if $T^* > 0$. Change is regressive if $T^* < 0$. Since it is hard to pinpoint when banks adopt those technologies and when the effects on their cost structure are fully realized, such a general specification of technological change has the advantage of not requiring the exact dating of those changes.

[12] see Baumol and Willig (1988)

Second, if technological changes affect banks differently over time, then there should be time-varying effects in the bank-specific terms. Such time-variant effects will affect the measure of efficiency. To illustrate, let us redefine equation (2) as,

$$y_{it} = X_{it}'\beta + \xi_i + u_{it} \tag{5}$$

where $\xi_i = \alpha_0 + \varepsilon_i$ and u_{it}, ε_i are white noise. For the cost function $X_{it} = [\log(W_j/W_1), (W_l W_j - 0.5 * (W_l^2 + W_j^2)), q_l \log(W_j/W_1), q_l, G_l]'$, and $\beta = [\beta_j, \gamma_{lj}, \delta_l, \alpha_l, \theta_l]'$ for each bank i.

Assuming time-invariant inefficiency, the cost frontier intercept is given by the minimum of the firm-specific effects, ξ_i, and inefficiency is the difference between the frontier, or the "best-practice" firm, and the firm effects. Such deviation from the "best-practice" firm can come from differences in management skills, human inertia, and adoption of technology.

$$\hat{\alpha}_0 = \min_j(\hat{\xi}_j) \quad \textit{and} \quad \hat{\varepsilon}_i = \hat{\xi}_i - \hat{\alpha}_0$$

Assuming that cost efficiency is constant over time is relatively implausible given the long time dimension of the panel. We therefore consider the following approach to modelling time-varying cost efficiency. Consider a fixed effects approach. For time-varying firm inefficiency,

$$\hat{u}_{it} = \Omega_{i1} + \Omega_{i2} t + \Omega_{i3} t^2$$

and the time-varying fixed effect is

$$\hat{\xi}_{it} = \hat{\Omega}_{i1} + \hat{\Omega}_{i2} t + \hat{\Omega}_{i3} t^2$$

This specification allows cost efficiency to be time-varying as well as different for each bank,

$$\hat{\alpha}_t = \min_j(\hat{\xi}_{jt}) \quad \textit{and} \quad \hat{\varepsilon}_{it} = \hat{\xi}_{it} - \hat{\alpha}_t$$

Cost efficiency is derived as

$$CE_{it} = \exp\{-\hat{\varepsilon}_{it}\}$$

Measures of cost efficiency allow us to rank the banks over time from the most to the least cost-efficient. The most efficient bank has a measure of cost efficiency equal to one and less efficient banks have measures below one. Measures of cost efficiency are consistent as $T \to \infty$ if the model is estimated by maximum likelihood and consistent if $T \to \infty$ and $I \to \infty$ if estimated by generalized least squares. The particular estimators are discussed in section (IV).

TO DO: To be complete I should include confidence intervals.

IV. Methodology

In this section we present the techniques used to estimate the cost function. Different estimation techniques are applicable depending on the assumptions we are willing to make.[13] The unique panel that we have consists of six banks and eighty-three quarterly observations. The availability of a long time-series and short cross-section leads to several natural parametric estimators.

A. Fixed Effects Model

Recall equation (5), a generic unobserved effects model.

$$y_{it} = X_{it}'\beta + \xi_i + u_{it}$$

[13] We acknowledge McIntosh (2002) point that nonstationarity may affect the consistency of all our estimates. That said, Phillips and Moon (1999) point out that in cases where the time-series is nonstationary it may still be possible to attentuate the strong noise by pooling the two dimensions. The interpretation of the coefficients would then be of long-run relationships among variables.

The fixed effects model assumes that we can capture differences across banks in the constant term, ξ_i. These effects can be correlated with X_{it}. However, we do need $E[u_{it}|X_{it},\xi_i]=0$. An alternative assumption would be to assume ξ_i comes from some distribution. However, in our case the cross-section draws are not random and it would be inappropriate to estimate the model using a random effects estimator.[14] The fixed effect estimator of β and the individual effect $(\xi_i+\mu_{it})$ are consistent as $T\rightarrow\infty$. The large-T component also allows us to evaluate the cost inefficiency apart from the statistical noise.

B. Stochastic Frontier Model

Next, we adopt the stochastic frontier approach to measuring the inefficiency level of Canadian banks. The stochastic frontier methodology assumes the fixed effect in equation (5) is decomposed into a constant α_0 and a firm-specific inefficiency variable, ε_i. This framework allows us to calculate the efficiency level of each bank relative to the"best practice" firm in the sample, rather than to the absolutely efficient firm. Inefficiencies are assumed to follow the truncated normal distribution, while the random errors follow a standard normal distribution. The logic is that the inefficiencies must have a truncated distribution because inefficiencies cannot be negative. Both the inefficiencies and the errors are assumed to be orthogonal to the explanatory variables in the model. Note that relative to the fixed effects estimator, the truncated normal assumption may be relatively inflexible and presumes that most firms are clustered near full efficiency.[15] We estimate this model by the maximum likelihood estimator (MLE), which is consistent as $T\rightarrow\infty$, assuming the following:

- $u_{it}\sim iidN(0,\sigma_u^2)$
- $\xi_i\sim iidN^+(\mu,\sigma_\xi^2)$
- u_{it} and ξ_i are independently distributed.

[14] For completeness we did estimate the model assuming random effects. The Breusch and Pagan Lagrange multiplier test for random effects rejects that assumption

[15] An alternative would be to adopt the thick frontier approach, where deviations from the predicted performance between the highest and lowest quartiles represent inefficiencies. One advantage of this approach is that it is distribution-free, but we do not have enough banks to construct such a thick frontier.

The log likelihood we maximize is given by[16]

$$\log L = -\frac{I(T-1)}{2}\log\sigma_u^2 - \frac{I}{2}\log(\sigma_u^2 + T\sigma_\xi^2) + \sum_i \log\left[1 - \Phi\Big(-\frac{\mu_i^*}{\sigma^*}\Big)\right]$$

$$-\Big(\frac{e'e}{2\sigma_u^2}\Big) + \frac{1}{2}\sum_i \Big(\frac{\mu_i^*}{\sigma^*}\Big)^2$$

where $\mu_i^* = \frac{T\sigma_\xi \bar{e_i}}{(\sigma_u^2 + T\sigma_\xi^2)}$, $\sigma^* = \Big(\frac{\sigma_\xi^2\sigma_u^2}{(\sigma_u^2 + T\sigma_\xi^2)}\Big)^{1/2}$, and $\bar{e_i} = (1/T)\sum_i(\xi_i + u_{it})$.

C. Seemingly Uncorrelated Regressions (SUR)

Thus far we have estimated the single equation cost function, ignoring the share equations, (3). Including the share equations allows us to increase the information without losing degrees of freedom. The system of equations will be estimated using maximum likelihood. Consider the following system of equations

$$\begin{aligned} C_{it} &= A_{it}\Theta + \xi_i, + u_{it} \\ S_{Kit} &= B_{it}\Gamma + \eta_{Ki} + \nu_{Kit} \\ S_{Lit} &= D_{it}\Phi + \eta_{Li} + \nu_{Lit} \quad i = 1,...,N \quad t = 1,...,T \end{aligned} \tag{6}$$

If we believe that shocks to Canadian banks are uncorrelated with each other, we can estimate separate models for each bank, using least squares. Equations (6)would be estimated using subscript t for each of the six banks. The result is six equations each with a large number of variables estimated using 82 periods. Results are not presented because of major problems with multicollinearity. This is because there is little variation in the regressors across time.[17]

However, if we think that disturbances are contemporaneously correlated across banks a more efficient estimator is SUR (Zellner (1962)). The SUR model encompasses least squares and therefore the assumption of no contemporaneous correlation can be tested. The assump-

[16] See Kumbhakar (2002) for details

[17] Note that this is even more severe if we include σ_{ij}, the cross-product output terms. The problem is however substantially improved when the equations of six banks are pooled.

tion of zero contemporaneous correlation is strongly rejected. Adding the cross-sectional component also allows us to identify the relevant parameters. Note that we iterate Zellner's estimator until it converges to the maximum likelihood estimator.

V. Data

The data set used in this paper consists of quarterly observations of the six largest banks in Canada starting from 1983q1 to 2003q3.[18] The data set came from the chartered banks' consolidated monthly balance sheet and quarterly consolidated statement of income collected by the Office of the Superintendent of Financial Institutions Canada (OSFI). The consolidated monthly balance sheet data at the aggregate level are published in Tables C1 and C2 in the Bank of Canada Banking and Financial Statistics. Large categories of the consolidated statement of income at the aggregate level are available in Table K2 in the same publication with an annual frequency.[19] All balance sheet data are end-of-month values and are converted to quarterly series by taking the quarterly average. The data set used in this paper is quarterly and deflated by the GDP deflator (1997=100). Appendix A provides a detailed definition of all variables used in this study, their summary statistics, and key financial statement ratios of the banks included in our sample.

Input prices

Since the definition of a bank's inputs and outputs is an on-going debate we choose to We follow the intermediation approach.[20] In the intermediation approach a bank is assumed to use labour, capital and deposits to produce earning assets. This is the approach commonly used in the conventional cost function literature. Here, deposits are treated as an input. The

[18] Our choice of the number of banks is limited by data availability. The six banks are: Royal Bank Financial Group, Bank of Montreal, Canadian Imperial Bank of Commerce, TD Bank Financial Group, Bank of Nova Scotia and National Bank. They are the only banks for which data are available in the entire sample period. The majority of the rest of the domestic chartered banks did not start reporting until after 2000. The big six account for around 90 per cent of the Canadian banking industry in terms of total assets.

[19] Disaggregate data are confidential.

[20] See Wang (2003) for a review.

production approach, however, postulates that banks also provide value-added in their deposit services. Deposits are treated as an output under the production approach.

There are three input prices in our model. L is the hourly wage of a bank's full-time equivalent employees. K is capital cost, measured by the expense on premises, and computer and equipment divided by the total stock of premises and fixed assets on the bank's balance sheet. D is the price of deposit, measured by the total interest expense on total deposits divided by total deposits. See table (XII) for summary statistics.

Output quantity

We define five output categories: Y1, consumer loans; Y2, non-mortgage loans; Y3, mortgage loans; Y4, security investment and Y5, non-traditional banking activities. The first four categories are taken from the asset side of a bank's balance sheet.[21] The measure of non-traditional activities is discussed below.

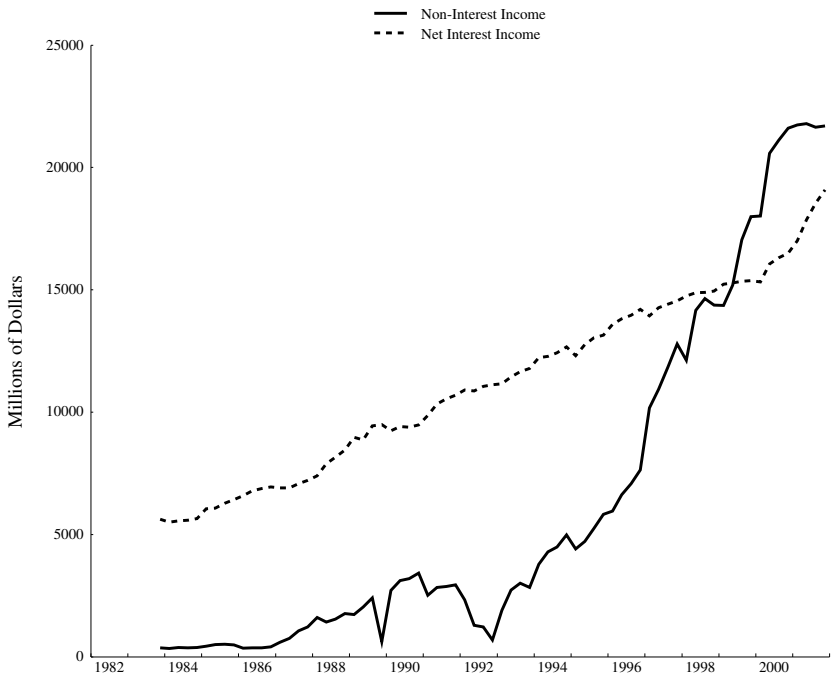
Non-traditional activities

Non-traditional banking activities are often overlooked in the literature of bank efficiency, and scale and scope economies. They are non-interest related activities such as depositor services, underwriting, foreign exchange trading and wealth management. Figure (1) shows that noninterest income has increased substantially since the late 1980s and has exceeded net interest income. The chart shows that banks have been shifting away from traditional lending and investment activities to non-traditional activities, a trend that is observed in U.S. banks Stiroh (2000).

Nontraditional activities can be divided into three categories according to their underlying asset or liability. The first type is generated from on-balance-sheet assets that are already captured in the output measures mentioned above. Examples include mortgage loan fees, gains and losses from trading and investment activities, and insurance income. The second

[21] All assets are reported in book value, except securities that are categorized under the trading account.

Figure 1. Canadian Chartered Bank Total Net Interest Income vs. Noninterest Income



type is generated from the liability-side of the balance sheet, such as deposit account fees and payroll processing fees. This type of output is often omitted in studies using an intermediary approach (as in our case), in which deposits are treated as inputs into the production of a bank. The third type comes from off-balance-sheet activities like securities underwriting, wealth management, the provision of loan guarantees and letters of credit. These activities are also ignored under the conventional approach to measuring bank output.

Rogers (1998) proposes to directly include noninterest income in the cost function as a measure of these activities. However, the resulting cost function mixes stocks (asset) and flows (revenue), for which the author provides no explanation. Boyd and Gertler (1994) introduces an asset-equivalent measure of these non-traditional activities.[22] Assuming that these

[22]The authors call it an asset-equivalent measure of off-balance-sheet activities, assuming that all noninterest income is generated from off-balance-sheet assets.

non-traditional activities yield the same rate of return on assets (ROA) as traditional activities, the assets that are required to produce noninterest income can be calculated by dividing noninterest income by the ROA of traditional activities. Their definition of bank profits is as follows:

$$\pi = II - IE - PROV - NE + NII$$

where II is total interest income; IE is total interest expense; $PROV$ is loan loss provisioning; NE is noninterest expense, and NII is noninterest income. Boyd and Gertler (1994) then assume that (1) noninterest income is generated by some hypothetical off-balance-sheet asset A_o and that (2) using the same capital and liabilities, A_o generates the same rate of return as on-balance-sheet asset A_b. The authors show that,

$$AEM = A_o = A_b[NII/II - IE - PROV]$$

AEM is called the asset-equivalent measure of non-traditional activities. This measure can be viewed as the hypothetical asset holdings that would be required to generate noninterest income. It has the convenience of using readily available data and it is easy to communicate. However, it has a few limitations. First of all, it is likely to be an overestimate of bank output that is orthogonal to what has been captured by other measures of output. As mentioned, some components of noninterest income are also generated from on-balance sheet assets. Ideally, we would like to be able to subtract those components from noninterest income. Unfortunately, disaggregate data of noninterest income are only available from 1997 onwards. [23] We consider this the best available proxy for such activities. Secondly, the assumption that off-balance-sheet assets yield the same rate of return as on-balance-sheet assets ignores the risk aspect of the two types of assets. It is also possible that if banks can achieve diversification effects, the two returns could be negatively correlated.[24]

Regulatory Changes

[23] Noninterest income generated from on-balance-sheet assets account for approximately 20 per cent of total noninterest income from 1997 to 2003.

[24] Such an effect, however, has been shown by Calmès and Liu (2003) to be minimal.

A bank's cost function can also be influenced by exogenous factors, such as changes in the regulatory environment. However, we should note that regulatory changes may not have been initiated aiming at reducing costs or improving efficiency. Rather, they are often the product of intermingling forces, such as evolutions in technological advances, demographic changes and global trends. However, by changing the regulatory environment in which banks operate, such changes may have helped or deterred banks in their cost-minimizing/profit-maximizing objectives.

Three notable changes to the Bank Act took place in our sample period. The most influential one occurred in 1987, when Canadian banks were permitted to invest in corporate securities, as well as distribute government bonds. Furthermore, banks are now allowed to purchase control of investment dealers and invest in the securities business. As a result, banks have substantially increased their financial-market based activities. On the demand side, as bank customers began to be able to invest in the financial market directly through their banks, the amount of direct financing (for example, financing done through financial markets as opposed to through financial intermediaries) also increased.

The situation evolved further when banks were given the right in 1992 to enter the trust business through the establishment or acquisition of trust companies in Canada. Over the next few years, the major banks bought up most of the trust companies. They were also allowed to offer a number of in-house activities such as portfolio management and investment advice. These changes may have attracted a larger fraction of depositors to invest in financial markets directly through their banks. Finally, various changes to update and refine the amendments made in 1992 took place in 1997.

Figure (1) shows a gradual take-off of the banks' noninterest income in 1987, reflecting a trend towards more non-loan-oriented activities. This trend grew further after the amendment in 1992 and continued throughout the 1990s.

Besides regulatory changes that allow banks to diversify their business mix, developments in legal reserve or capital requirements aimed at ensuring the financial soundness of banks may have affected the banks' cost structure. In 1989, Canadian banks adopted the minimum

capital requirement proposed by the Basel Committee on Banking Supervision.[25] This may have at times affected a bank's output decisions if capital requirement is binding. In addition, the removal of legal reserve requirement in 1991 and the complete phasing out of reserve requirement by banks in 1994 could also have an impact on the banks' production decisions.

To investigate whether such structural changes in a bank's revenue source would have a significant impact on a bank's cost structure, we include three regulatory dummies, 1987Q2, 1989Q1, 1991Q1, 1992Q1, 1994Q1 and 1997Q1, in our cost function. The dummies are equal to zero before these dates and equal to one afterwards.

VI. Results

In this section we present our findings. Equations (2) and (4) are estimated by SUR. Equation (2) is also estimated by fixed effects-GLS and by single-equation MLE. For each method we estimate two models. The first model, labelled Model REG, includes important regulatory changes. The second model, labelled Model T (for time dummies), includes time dummies that are assumed to proxy technological change.[26] Parameter estimates of each model are presented in the appendix.

A. Diagnostic Tests

Before we discuss the results of economies of scale and scope and relative efficiency, we conduct two diagnostic tests. The first one is on the own- and cross-price elasticity of the cost function. Table (I) presents the own- and cross-price elasticities for the SUR specification. All own-price elasticities are negative and their absolute values are less than one, implying that all inputs are price-inelastic. Capital has the highest own-price elasticity among the three inputs. This is interesting because capital expense makes up the smallest share of input (see table (XIV) in the appendix). Deposits, by far the largest component in the input mix, have

[25] Banks are required to hold a minimum of 8% capital relative to risk-weighted assets.

[26] When both types of dummies are included in the model, the cost function is non-concave everywhere.

the lowest price sensitivity. This is reasonable as banks are likely to shift any price changes in its sources of funding to their customers by charging a higher rate of interest on loans or demanding a higher rate of return on its securities. However, changes in capital expenses may be much harder to absorb. The cross-price elasticities between labour and capital, and between labour and deposits are roughly the same, suggesting similar substitutability between the two pairs of inputs. It is puzzling, however, the cross-price elasticity between capital and deposits is negative. Note that this elasticity measure is statistically insignificant.

Table I
Own- and Cross-Price Elasticities

	Capital	Labor	Deposits
Capital	-0.8709 (0.0159)		
Labor	0.6862 (0.1241)	-0.8000 (0.0970)	
Deposits	-0.2746 (0.2862)	0.7709 (0.1442)	-0.3391 (0.1233)

Standard deviations are in parentheses

The second test is for the validity of restrictions on the production structure. Table (II) shows the results on testing an unrestricted cost function against those with restrictions to ensure homothenticity and homogeneity.[27] The production function is homothetic if $\gamma_{lj} = 0, \forall\, l, j$. Furthermore, the production function is homogeneous of degree one if $\sum_l \alpha_l = 1$. Restrictions are strongly rejected. That is, a homothetic or homogenous production structure for the Canadian banking sector is statistically rejected.[28]

B. Economies of Scale and Scope

Results on economies of scale are presented in Table (III). We test the null hypothesis of constant returns to scale (CRS) and present test statistics and the p-values of the test. The likelihood ratio test is distributed chi-squared with one degree of freedom. Overall, the results from all six models are in favour of increasing returns to scale. All six economies of scale

[27] If we restrict the model to an aggregate output we could test the Cobb-Douglas restriction. However, in our sample banks are clearly multi-product institutions.

[28] We only report results for the stochastic frontier model. Results for the fixed effects model are similar.

Table II
Tests of Production Structure

Model	Log Likelihood	Test Statistic	Degrees of Freedom	P-value
Unrestricted model				
Model REG	999.56		31	
Model T	1021.96		31	
Homothenicity				
Model REG	886.31	226.38	16	0.0000
Model T	901.21	241.50	16	0.0000
Homogeneity				
Model REG	978.95	41.22	24	0.0000
Model T	990.28	63.36	24	0.0000

measures are greater than one and statistically significant. The economies of scale measures are smaller under Model T than under Model REG. Evaluated at the sample mean, the measured economies of scale under Model T range between six per cent and 13 per cent. That is, a 1 per cent increase in output will raise production cost by 0.87 per cent to 0.94 per cent, depending on the methodology. The implied cost savings are even higher under Model REG, ranging between 18 to 20 per cent. This magnitude is stronger than those typically found in the literature on large banks. For example, Stiroh (2000) finds significant economies of scale in medium to large sized US Banking Holding Companies. In their study, the average measure of economies of scale for banks with assets of more than five billion dollars in 1991-1997 is around five per cent.[29]

As a robust check, we also evaluate the economies of scale for the sub-periods 1983-1986, 1987-1996 and 1997-2003, to reflect changes in the Bank Act that are found to be significant in our model. As shown in tables (XV) and (XVI) in appendix B, all six models still find significant economies of scale in all three sub-periods. The magnitude of average cost savings associated with a higher output, however, evolves over time. It seems to be consistent across all models that the degree of economies of scale dropped in 1987-1996, and reached the highest

[29]The six banks in our study have an asset size of 80 to 400 billion dollars in 1991-1997.

in 1997-2003. In other words, the evolution of economies of scale resembles a U-shaped curve.[30]

Table III
Economies of Scale

Model	ζ	$H_0 : \zeta = 1$	
		Statistic	P-value
SUR			
Model REG	1.181	67.47	0.0000
Model T	1.064	9.48	0.0021
single-FEGLS			
Model REG	1.197	63.52	0.0000
Model T	1.126	19.73	0.0000
single-MLE:			
Model REG	1.176	50.08	0.0000
Model T	1.083	7.50	0.0062

[a]The restriction imposed on equation (2) is actually $\zeta^{-1} = 1$ and $\sum_j \delta_{lj} = 0 \; \forall \; l$ since returns to scale is defined as $\frac{\partial C}{\partial q_l} = \sum_l \alpha_l + \sum\sum \delta_{lj} \log(\bar{P}_j/\bar{P}_1)$ where $\bar{\cdot}$ is the sample mean

Table (IV) shows the measures of economies of scope derived from all three econometric specifications under Model T across pairs of bank outputs. The results between the fixed effect model and the stochastic frontier model are highly similar. Overall, the evidence of economies of scope is mixed. We find two pairs of outputs that exhibit significant economies of scope: personal and real estate loans, and real estate loans and non-traditional activities.[31] This finding that real estate loans have product complementarity with other products is similar to Murray and White (1983) who find significant product compatibility between mortgage lending and other lending for BC credit unions. However, we also find significant diseconomies of scope between personal and business loans, and between business and real estate loans. This may indicate a poor product complementarity between business loans and other

[30]As another robust check, we also try excluding non-traditional activities as a bank output, as most other studies do. The measured economies of scale is still significant but slightly smaller, ranging from three to 18 per cent, depending on the model.

[31]Our criterion is that the estimated economies of scope must have the right sign (negative for economies of scope and positive for diseconomies of scope) and significant in at least two out of the three models, and insignificant in the remaining model if the latter indicates an opposite sign.

business lines. The measures of economies of scope on other product combinations are inclusive. Our finding of weak evidence of economies of scope is consistent with the existing literature on U.S. banks.

Table IV
Economies of Scope: estimate and p-values

Variable	η_{ij}^{SUR}	η_{ij}^{FE}	η_{ij}^{MLE}
personal-business	0.099	4.294	4.342
	(0.0052)	(0.0095)	(0.0089)
personal-real estate	0.012	-1.265	-1.208
	(0.4567)	(0.0121)	(0.0119)
personal-securities	0.085	0.532	0.592
	(0.0858)	(0.5925)	(0.5444)
personal-OBS	0.041	-0.771	-0.839
	(0.0234)	(0.1447)	(0.1146)
business-real estate	-0.021	2.889	2.900
	(0.3427)	(0.0269)	(0.0246)
business-securities	-0.152	-1.215	-1.421
	(0.0000)	(0.5083)	(0.4405)
business-OBS	-0.073	1.760	2.015
	(0.0093)	(0.0639)	(0.0367)
real estate-securities	-0.018	0.358	0.395
	(0.0422)	(0.5096)	(0.4428)
real estate-OBS	-0.009	-0.519	-0.560
	(0.0322)	(0.0814)	(0.0334)
securities-OBS	-0.063	0.218	0.274
	(0.0005)	(0.6030)	(0.5480)

[b] P-values for the hypothesis that $\eta = 0$ are presented in parentheses

C. Technological and Regulatory Changes

Recall that our proxy for technological change is given by by

$$T^* = -\frac{\partial C}{\partial t} = -(\theta_1 + 2\theta_2)$$

Table V
Technological Change

Model	T^*	$H_0 : T^* = 0$	
		Statistic	P-value
SUR	0.01062	216.65	0.0000
FE-GLS	0.00934	83.72	0.0000
MLE	0.00975	89.76	0.0000

Table VI
Regulatory Change

Model	$H_0 : D = 0$			
	D_{1987}	P-value	D_{1997}	P-value
SUR	-0.0974	0.01	-0.0989	0.01
FE-GLS	-0.04903	0.01	-0.05208	0.01
MLE	-0.0491	0.01	-0.05591	0.01

Table (V) presents the estimates of T^* in Model T under our three methodologies. We find significant evidence of increasing cost-efficiency over the sample period in all three models.

We have included six regulatory dummies in our cost function: 1987, 1989, 1991, 1992, 1994 and 1997. As expected, the dummies for 1987 and 1997 are highly significant in the cost function, suggesting that regulatory changes to the Bank Act in those years to allow banks to be more diversified did have an impact on the banks cost structure. However, the other dummmies were not significant and thus were not included in the regressions. Table (VI) shows the estimates and significance values for the regulatory dummies in our three specifications. As expected, all models yield negative coefficients for the dummies, implying that regulatory changes led to a decrease in costs. Again, the results of our two favoured models are highly similar, suggesting roughly a five per cent cost savings following either regulatory change.

Table VII
Relative cost-efficiency

Bank	MLE	Fixed Effects	SUR
Bank 1	100	100	100
Bank 2	98.1	98	95.1
Bank 3	96	96.5	95
Bank 4	91.4	92.4	92
Bank 5	89.8	90.6	90.5
Bank 6	79.4	79.3	77.2

D. Relative Efficiency

In table (VII) we present measures of relative efficiency calculated from the fixed effects model and stochastic frontier model (MLE) for the big six Canadian banks, assuming relative efficiency is time-invariant. Bank 1 is assumed to be on the efficiency fronier, i.e. the most efficient bank among the six. Except for one bank, the measured inefficiency of all banks is within ten per cent, which is similar to those found the the literature on U.S. banks and bank holding companies.[32]

We also find that the relative ranking of efficiency between banks seems to be positively correlated with bank size.[33] Given that scale economies are already taken into account in our cost function, the remaining difference in bank specific effects may be related to other factors such as management skills and the adoption of new technology. This result suggests that a smaller bank may be able to benefit from economies of scale by simply increasing output, but this may not be sufficient to achieve the same efficiency of a larger bank.

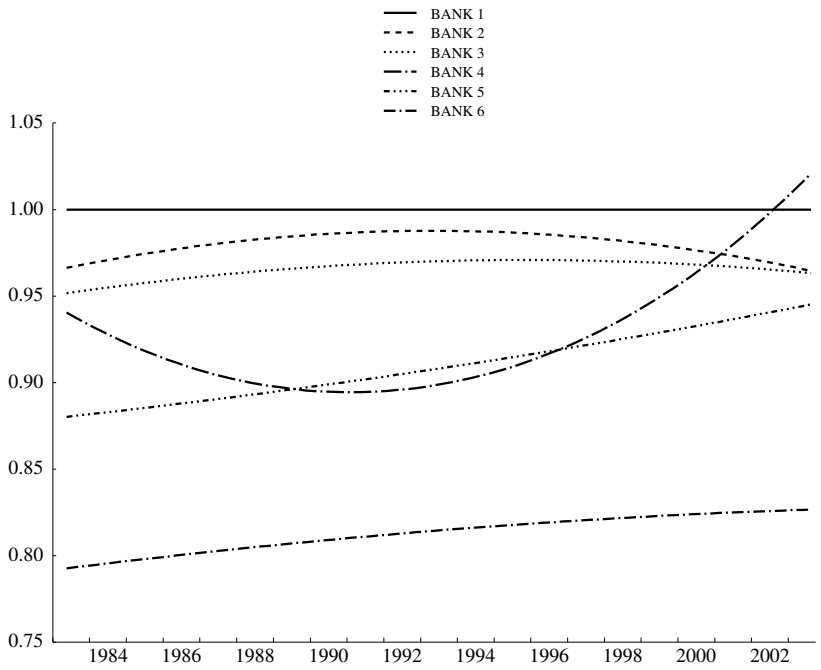
Our calculations of time-varying cost-efficiency from the stochastic frontier model are almost identical to those from the fixed firm-effect model. Figure (2) is the profile of the different relative firm-effects over time. Overall, there is little change in the ranking of relative efficiency in our sample period. The exception is Bank 4, whose efficiency ranking dropped from the fourth in the early to mid-1980s to the fifth in the early 1990s and then rose to the

[32]For example Berger and Mester (1999) and Stiroh (2000). The measure of profit inefficiency is usually higher, around 20 to 40 per cent.

[33]The identify of the Banks cannot be disclosed due to the sue of confidential data.

most efficient bank in 2002. The average of the measures of inefficiency is very similar to those from the time-invariant model, with most banks having inefficiency of roughly ten per cent over the sample and Bank 6 having 20 per cent. The dispersion of inefficiency between banks seems to have narrowed over time.

Figure 2. Relative Efficiency of Large Canadian Banks



VII. Conclusion

We have applied the translog cost function framework to study the efficiency, economies of scale and scope of the six largest Canadian banks. Using a unique panel dataset from 1983q1 to 2003q3, we estimate three econometric models based on this framework. Given the long time dimension of the dataset, we add a time trend and a time trend squared in the cost function to capture any technological change over time. We find that banks have experienced

significant technology progress over the past 20 years. We also find that regulatory changes have helped to reduce the production cost of banks

Overall, we reject the hypothesis of constant returns to scale. That is, banks can enjoy cost savings by increasing their scale of production. This finding is similar to recent studies on large U.S. banks, while our measured economies of scale is larger than those found in the U.S. studies. Our findings contrast those found by McKinsey & Company (1998), which fails to uncover economies of scale in Canadian banks.[34] The findings regarding economies of scope are mixed, with only economies of scope found in combinations between mortgage loans and some other outputs, and diseconomies of scope found between business loans and other activities.

Measures of efficiency are derived from the efficiency frontier estimated by the stochastic frontier model. On average, the inefficiency of Canadian banks is around ten per cent, close to what is typically found in the literature on U.S. banks using the cost function approach. The ranking of efficiency also suggests that larger banks seem to be more cost-efficient than smaller banks. Given that scale economies are already accounted for in our model, such heterogeneous effects may come from differences in other factors such as management skills and adoption of technology. Our time-varying fixed-effect panel specification allows us to trace the changes in efficiency levels over time. The results suggest that the relative efficiency level of banks has remained unchanged, and the dispersion between banks seems to have narrowed.

There are several possible extensions. We have taken an intermediation approach in modelling a bank's production, i.e. assets are treated as outputs and deposits are treated as an input rather than output. There has been on-going debate regarding what defines a bank's output. Wang (2003) proposed alternative measures of bank output based on banks' value added. The author has shown traditional measures of output such as the one we use overestimate bank output by more than 25 per cent. Wang also argues that non-traditional activities can be better modelled under the new measure.

[34] McKinsey & Company bases its analysis on the correlation between size and efficiency ratio (non-interest expense over total revenue), which says little about economies of scale.

VIII. Appendix

Table VIII: **SUR: Model REG**

Variable	**Equations**		
	Equation $\log(c)$	Equation S_k	Equation S_l
D_1	-0.09742** (0.01046)		
D_2	-0.09887** (0.01094)		
α_1	-0.23412** (0.08385)		
α_2	0.42090** (0.07851)		
α_3	0.05055 (0.05574)		
α_4	0.36156** (0.10125)		
α_5	0.17454** (0.05351)		
β_1		-0.24733** (0.03077)	
β_2	-1.21306** (0.06539)		-1.21306** (0.06539)
β_3	2.46039** (0.07454)		
δ_{12}	0.03101** (0.00171)	0.03101** (0.00171)	0.03101** (0.00171)

Continued on next page...

... table VIII continued

Variable	**Equations**		
δ_{13}	-0.04859** (0.00132)	-0.04859** (0.00132)	
δ_{23}	-0.12333** (0.00305)		-0.12333** (0.00305)
γ_{11}		0.00335 (0.00286)	
γ_{12}	0.03074** (0.00597)		0.03074** (0.00597)
γ_{13}	-0.03408** (0.00704)		
γ_{21}		-0.0102** (0.00261)	
γ_{22}	-0.00618 (0.0054)		-0.00618 (0.0054)
γ_{23}	0.01637** (0.00625)		
γ_{31}		0.01437** (0.00192)	
γ_{32}	0.00496 (0.004)		0.00496 (0.004)
γ_{33}	-0.01933** (0.00472)		
γ_{41}		0.0036 (0.00352)	
γ_{42}	-0.01254† (0.00725)		-0.01254† (0.00725)

Continued on next page...

... table VIII continued

Variable	Equations	
γ_{43}	0.00894 (0.00839)	
γ_{51}	-0.01593** (0.0018)	
γ_{52}	-0.01085** (0.0037)	-0.01085** (0.0037)
γ_{53}	0.02678** (0.00442)	
Log-likelihood	3656.7	

Table IX
Fixed Effects

	Model REG		Model T	
Variable	**Coefficient**	Std. Error	**Coefficient**	Std. Error
α_1	-1.37128*	(0.54683)	-1.09263*	(0.52506)
α_2	3.13151**	(0.4852)	2.59261**	(0.46146)
α_3	-0.92259*	(0.35976)	-1.28914**	(0.34680)
α_4	0.38785	(0.61813)	0.56971	(0.59401)
α_5	-0.56202†	(0.29758)	-0.18192	(0.28848)
β_2	4.20985**	(0.82177)	4.15805**	(0.79775)
β_3	-0.20342	(0.30928)	-0.65401*	(0.29332)
δ_{12}	0.33198**	(0.05356)	0.33826**	(0.05099)
δ_{13}	-0.17716**	(0.02256)	-0.18761**	(0.02154)
δ_{23}	0.03414	(0.02387)	0.04373†	(0.02266)
γ_{12}	0.12243**	(0.0456)	0.09445*	(0.04385)
γ_{13}	-0.04647†	(0.02543)	-0.07026**	(0.02379)
γ_{22}	-0.21692**	(0.0404)	-0.18215**	(0.03829)
γ_{23}	0.14395**	(0.01901)	0.11644**	(0.01948)
γ_{32}	0.07832**	(0.02987)	0.12281**	(0.02894)
γ_{33}	-0.09132**	(0.01827)	-0.07344**	(0.01781)
γ_{42}	-0.0243	(0.0517)	-0.03226	(0.04965)
γ_{43}	-0.04092	(0.02559)	0.00149	(0.0248)
γ_{52}	0.06105*	(0.02389)	0.03021	(0.0232)
γ_{53}	0.06413**	(0.01625)	0.06698**	(0.01552)
α_0	-20.84538*	(8.50474)	-20.92548*	(8.20029)
D1	-0.04903**	(0.01065)		
D2	-0.05208**	(0.01087)		
θ_1			-0.00944**	(0.00102)
θ_2			0.00005**	(0.00001)

N	498	498
R^2	0.9907	0.9914
$F_{(27,470)}$	2265.9	2464.4

Table X
Stochastic Frontier Model

	Model REG		Model T	
Variable	**Coefficient**	Std. Error	**Coefficient**	Std. Error
α_1	-1.3446*	(0.53479)	-1.01653*	(0.51572)
α_2	3.22947**	(0.47598)	2.73162**	(0.45644)
α_3	-0.89810*	(0.35182)	-1.31962**	(0.33982)
α_4	0.44000	(0.60459)	0.59944	(0.58132)
α_5	-0.62384*	(0.29171)	-0.21008	(0.2821)
β_2	4.25144**	(0.80359)	4.12266**	(0.77941)
β_3	-0.23031	(0.30256)	-0.70444*	(0.28788)
δ_{12}	0.31716**	(0.05271)	0.31407**	(0.05082)
δ_{13}	-0.17143**	(0.02217)	-0.17941**	(0.02136)
δ_{23}	0.03165	(0.02336)	0.03974†	(0.02226)
γ_{12}	0.12183**	(0.04459)	0.09109*	(0.04299)
γ_{13}	-0.03984	(0.02499)	-0.05723*	(0.02401)
γ_{22}	-0.22353**	(0.03958)	-0.19136**	(0.03774)
γ_{23}	0.14899**	(0.0187)	0.12593**	(0.0194)
γ_{32}	0.07591**	(0.02922)	0.12573**	(0.02836)
γ_{33}	-0.09451**	(0.01791)	-0.08028**	(0.0177)
γ_{42}	-0.02919	(0.05057)	-0.03567	(0.04858)
γ_{43}	-0.04268†	(0.02503)	-0.00366	(0.0244)
γ_{52}	0.06525**	(0.02339)	0.03103	(0.02269)
γ_{53}	0.05976**	(0.01599)	0.0605**	(0.0153)
α_0	-23.08205**	(8.34915)	-23.49573**	(8.10225)
D1	-0.0491**	(0.01042)		
D2	-0.05591**	(0.01075)		
θ_1			-0.00987**	(0.001)
θ_2			0.00005**	(0.00001)
N	498		498	
Log-likelihood	938		961.3	
$\chi^2_{(22)}$	51834		55467	

References

Baumol, W.J., Panzra J.C., and R.D. Willig, 1988, *Contestable Markets and the Theory of Industry Structure*. (Harcourt Brace Jovanovich) second edn.

Berger, Allan N., and David B. Humphrey, 1997, Efficiency of Financial Institutions: International Survey and Directions for Future Research, *European Journal of Operational Research* 98, 175–212.

Berger, Allan N., and Loretta J. Mester, 1999, hat Explains the Dramaticv Changes in Cost and Profit Performance of the U.S. Banking Industry?, Wharton School Center for Financial Institutions, University of Pennsylvania, working paper 99-10.

Berger, A.N., G.A.Hanweck, and D.B.Humphrey, 1985, Competitive Viability in Banking: Scope and Product Mix Economies, working paper, Division of Statistics, Board of Governors of the Federal Reserve System.

Boyd, J., and M. Gertler, 1994, Are Banks Dead? Or Are the Reports Greatly Exaggerated, Summer 18(3) Federal Reserve Bank of Minneapolis Quarterly Review.

Breslaw, Jon A., and James McIntosh, 1997, Scale Efficiency in Canadian Trust Companies, *Journal of Productivity Analysis* 8, 281–292.

Calmès, C., and Y. Liu, 2003, Financial Structure Change and Banking Riskiness: the Case of Canada, Bank of Canada Working Paper, forthcoming.

Christensen, L.R., D.W.Jorgenson, and L.J. Lau, 1971, Conjugate Duality and the Transcendental Logarithmic Production Function, *Econometrica* 39, 255–256.

Clark, Jeffrey A., and Thomas F. Siems, 2002, X-Efficiency in Banking: Looking beyond the Balance Sheet, *Journal of Money, Credit, and Banking* 34, 987–1013.

Cornwell, C., P. Schmidt, and R.C. Sickles, 1997, Production Frontiers with Cross-sectional and Time-series Variation in Efficiency Levels, *Journal of Econometrics* 46, 185–200.

Diewert, W.E., and R.J.Kopp, 1982, The Decomposition of Frontier Cost Function Deviations into Measures of Technical and Allocative Efficiency, *Journal of Econometrics* 19, 319–331.

Ferrier, G.D., and C.A.K.Lovell, 1990, Measuring Cost Efficiency in Banking, *Journal of Econometrics* 46, 229–245.

Freedman, C, and C. Goodlet, 1998, The Financial Services Sector: Past Changes and Future Prospects, Bank of Canada Technical Report 82.

Kumbhakar, S.C, 2002, *Stochastic Frontier Analysis*. (Cambridge University Press) first edn.

Lawrence, Colin, 1989, Banking Costs, Generalized Functional Forms, and Estimation of Economies of Scale and Scope, *Journal of Money, Credit and Banking* 21, 368–379.

McIntosh, James, 2002, A Welfare Analysis of Canadian Chartered Bank Mergers, *Canadian Journal of Economics* 35, 457–475.

Mitchell, K., and N.M. Onvural, 1996, Economies of Scale and Scope at Large Commercial Banks: Evidence from the Fourier Flexible Functional Form, *Journal of Money, Credit and Banking* 28, 178–199.

Murray, John D., and Robert W. White, 1983, Economies of Scale and Economies of Scope in Multiproduct Financial Institutions: A Study of British Columbia Credit Unions, *Journal of Finance* 3, 887–902.

Rezvanian, Rasoul, and Seyed Mehdian, 2002, An Examination of Cost Structure and Production Performance of Commercial Banks in Singapore, *Journal of Banking and Finance* 26, 78–98.

Rime, Bertrand, and Kevin J. Stiroh, 2003, The Performance of Universal Banks: Evidence from Switzerland, *Journal of Banking and Finance, forthcoming*.

Rogers, Kevin E, 1998, Nontraditional Activities and the Efficiency of US Commercial Banks, *Journal of Banking and Finance* 22, 467–482.

Schmidt, P., and C.A.K. Lovell, 1979, Estimating Technical and Allocative Inefficiency Relative to Stochastic Production and Cost Frontier, *Journal of Econometrics* 9, 343–366.

Shaffer, S., 1993, A Test of Competition in Canadian Banking, *Journal of Money, Credit, and Banking* 25, 49–61.

Stiroh, Kevin J, 2000, How Did Bank Holding Companies Prosper in the (1990)s?, *Journal of Banking and Finance* 24, 1703–1747.

Wang, J. Christina, 2003, Loanable Funds, Risk, and Bank Service Output, Federal Reserve Bank of Boston, Working Paper.

Zellner, A, 1962, An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias, *Journal of the American Statistical Association* 57, 348–368.

Appendix A. Data Description

Table XI
Data Description

		Variable Definitions
Y1	Consumer Loans	Dollar value of personal loans for non-business purposes
Y2	Non-mortgage loans	Dollar value of secured call and other loans to investment dealers and brokers + loans to regulated financial institutions + loans to domestic and foreign governments + lease receivables + reverse repurchase agreements + loans to individuals and others for business purposes
Y3	Mortgage loans	Dollar value of residential and non-residential mortgage loans
Y4	Other	Dollar value of other financial assets on a bank's balance sheet
Y5	OBS	Asset-equivalent measure of off-balance-sheet activties
L	Price of Labor	Total salaries, pensions and other staff benefits divided by the number of full-time equivalent employees and hours in a year
K	Price of capital	Rental expense on real estate and depreciation on premises, furniture, fixture, computer and equipment divided by total stock of land, buildings, and equipment, less accumulated depreciation
D	Price of deposits	Total interest expense on deposits divided by the total dollar amount of deposits
C	Variable costs	interest cost + labour expenses + capital costs

Table XII
Statistics of Real Input Price Growth

Variable	Mean	Standard Deviation
Hourly Wage	1.05	7.02
Unit Price of Capital Expense	0.95	10.85
Unit Price of Deposits	-1.88	9.92

[c]Units are quarter-over-quarter percentages

Table XIII
Mean of Data and Percentage in Assets

Variable	Billions of dollars	Percentage of asset
Total asset	798	100
Securities	297	37.2
Non-mortgage loans	260	32.6
Mortgage loans	148	18.6
Personal loans	76	9.6
Fixed asset	7	0.9
Deposits	564	70.7
Equity	36	4.6

[d]Calculated at sample mean, percentages do not necessarily sum to 100.

Table XIV
Key Balance Sheet Ratios

Variable	Percentage
Labour expense / noninterest expense	55.8
Capital expense / noninterest expense	18.8
Interest expense / noninterest expense	174.0
Noninterest expense / total revenue	42.0

Further Results on Economies of Scale

Table XV
Economies of Scale: REG Model

Model	ζ	$H_0 : \zeta = 1$	
		Statistic	P-value
1983 - 1986	1.1820	67.47	0.0000
1987 - 1996	1.1821	67.36	0.0000
1997 - 2003	1.1792	66.64	0.0000
single-FEGLS			
1983 - 1986	1.2015	58.96	0.0000
1987 - 1996	1.1749	44.30	0.0000
1997 - 2003	1.2154	72.91	0.0000
single-MLE:			
1983 - 1986	1.1779	45.48	0.0000
1987 - 1996	1.1519	33.07	0.0000
1997 - 2003	1.1959	60.79	0.0000

Table XVI
Economies of Scale: T Model

Model	ζ	$H_0 : \zeta = 1$	
		Statistic	P-value
SUR			
1983 - 1986	1.0636	9.46	0.0021
1987 - 1996	1.0596	8.13	0.0043
1997 - 2003	1.0666	10.60	0.0011
single-FEGLS			
1983 - 1986	1.1323	20.74	0.0000
1987 - 1996	1.0990	11.54	0.0007
1997 - 2003	1.1486	25.70	0.0000
single-MLE:			
1983 - 1986	1.0860	7.65	0.0057
1987 - 1996	1.0540	3.00	0.0831
1997 - 2003	1.1063	12.21	0.0005