Does information technology provide banks with profit?

Wesley Shu a,*, Paul A. Strassmann b

a Department of Information and Decision Systems, San Diego State University, San Diego, CA 92182-8234, USA
b Strassmann Inc., New Canaan, CT, USA

Accepted 13 June 2003
Available online 19 February 2005

Abstract

While many studies have affirmed the contributions of information technology (IT) to business value, people are not convinced. So far IT in the service industry has not yet been seen to be more productive. The data in most previous studies either focus on specific industries or exclude financial industry data. As such, the need to do an analysis on IT productivity in the service industry is imminent. We chose 12 banks covering 9 years for our analysis. To eliminate possible estimation errors, we applied an analysis for panel data—a random effect model. We found IT investment demonstrated the highest marginal product among the input factors we chose.

# 2005 Elsevier B.V. All rights reserved.

Keywords: Productivity; Panel data; Information technology; Banking

1. Introduction

Information systems (IS) or information technology (IT) productivity has always been a concern in academia and industry. The Bureau of Labor Statistics and National Income and Product Account have shown that IT investment has increased to 25 times what it was 30 years ago. During the same period, labor productivity did not increase. The labor productivity growth rate declined from its high of 2.68%/year in 1960s to a low of 1.03% in the early 1990s. This phenomenon has been labeled the “IT productivity paradox.” Nobel Laureate Solow’s famous saying succinctly pictured this paradox: “you can see the computer age everywhere but not in the productivity statistics.” [24] Some studies [4,5,19–21] also reported non-significant or negative IT contributions to business value. The absence of any positive correlation between profitability of firms and IT spending has been demonstrated by Strassmann, based on his consulting practice for 40 corporate cases [25], as well as for 292 corporate cases [26] and again for 486 corporate cases [27].

On the other hand, some research [6–8,18,23,29] found very positive IT contributions. After observing many studies showing positive IT contributions to business value, Bakos raised another question, “how can computers be so productive?” [1] However, by
using the same data set as the above five studies, Lee and Shu could not find significant IT contributions either. They also showed that difference in methodology could lead to a significant difference in research outcome [30] and also reported mixed results for newly industrialized economies.

While much research has been performed in the firm or at the macroeconomic levels, we have not seen many scholarly articles that analyze a specific industry. While some previous studies analyzed IT productivity in the manufacturing industry using data from the 1980s, this paper is focused on the banking industry using more recent data. In addition, our data came from a pooled data set, which contains both cross-section and time series data. Traditional ordinary least square (OLS) analysis suffers from a variety of statistical problems and does not produce satisfactory results; therefore some analysis tools for panel data, which can avoid these statistical problems, were used.

We have seen some studies of the productivity of capital, but little has been done to communicate the productivity of corporate information-creating and -consuming resources in ways that is useful to business executives in assessing, planning and budgeting. Thus, we resorted to IT budget information from CIOs to calculate IT capital.

2. IT and banking industry

Even though IT has become an omnipresent term in today’s business world, its definition is not clear. The U.S. Department of Commerce’s Bureau of Economic Analysis (BEA) has a category called “producers’ durable equipment—information processing equipment.” It includes office, computing and accounting machinery, computers and peripheral equipment, instruments, photocopy, and related equipment. However, the BEA does not include corporate operating expenses associated with IT budgets, such as the costs of administrators, systems planners, consultants, and equipment operators in a broader category.

Using government data, such discrepancies make IT productivity analysis unreliable. Thus, researchers have used other sets of IT spending data, for example, the data came from the International Data Group (IGD) survey on about 300 companies. But, the reliability of such a data set is still questionable because it used mail-in questionnaires or telephone surveys which are either incomplete or from interpretations that deal more with the views of the respondents than the facts. The respondents are usually administrators at the non-supervisory level.

We chose the banking industry because it is part of the service industry that has been suspected of having one of the lowest IT productivity. When Hitt and Brynjolfsson claimed that the IT productivity paradox was gone in their 1993 paper [12], many still believed that the service industry did not truly escape from this paradox. Ives [16] describes the argument as, “much of the hoopla surrounding the productivity paradox has centered on the high growth service industry.” Hitt and Bryjolfsson estimated IT had 81% gross marginal product (increase in dollar output per dollar of capital stock while the relative importance of computer capital assets was declining as a share of IT budgets) for their 1987–1991 data. The service industry still had only 0.7% productivity growth on average, far lower than the manufacturing industry average. Such a difference stirs our interest in knowing the cause. We are also interested to learn if such a suspicion is supported by empirical evidence and if recent data still demonstrate this paradox. Thus, the objective of this paper was to critically analyze the banking industry with more recent data and employ a model that fits better into the panel data that will be used.

3. Description of the data

Knowing the data source is necessary when assessing the reliability of the data used in IT productivity or profitability analysis. In our study, we use a proprietary data set obtained from 12 U.S. banks. The general financial data are taken from the Worldscope Global Researcher Database [9], which summarizes data filed with the Security Exchange Commission in a standardized format.

The IT data were obtained directly from the banking CIOs in connection with consulting services that analyzed their IT productivity. As IT spending, we employed the IT budgets of a bank’s central organizations as a reasonable approximation, because such data are not made public. Our data have higher reliability than that simply obtained from telephone or
mail-in surveys. Since the banks provided IT budget data for the practical purpose of understanding their productivity, this information must be a good proxy to IT spending from the banks’ own point of view. It is important, however, to note that this approach is not likely to suit most U.S. industrial corporations, where an increasing share of IT spending is absorbed as operating overhead expenses. In the banking industry, the concentration of IT spending under corporate management assures that there is a full accounting for all IT costs. In our case, we used the reported IT budget data and filled missing time series with imputed values. The imputation is based on a relationship between each bank’s IT budget and its non-interest expenses that represent the information-handling overhead’s costs.

Our data cover most of the largest banks in the U.S., including Bank One, Bank of America, Bank of Boston, Bankers Trust, Chase Manhattan, First Chicago, First Union, Fleet Financial, Keycorp, Morgan (J.P.) PNC, and Wells Fargo. The data spanned from 1989 to 1997 for each bank—a typical panel data set. Compared to previous studies, this data set is more up-to-date, reflecting recent conditions. Furthermore, since the data cover the 1990s, the statistical result provides an interesting comparison to the results from data of the 1980s.

Our input vector contains the IT budget, non-interest expenses, interest expenses, staff costs, and other operating expenses. Except for interest expenses, our input variables are IT and information management related. The reliance on non-interest expenses and other operating expenses is an indication of information management costs (e.g., overhead), which warrants particular attention in connection with our findings. It was noted by [28], that assessments of IT effectiveness must always start with an examination of the information productivity of industrial corporations. IT costs represent only 3–8% of the costs that a firm incurs for managing its information resources, which are represented mostly by the payrolls for administrative staffs and for purchases of information services. Before one can evaluate the effectiveness of IT one must first consider the productivity of the people who are using the IT. In this view of productivity, only people (with or without IT) can be productive because computers, by themselves, are only inert metal, glass, and plastics.

For industrial corporations, the cost of information resources is conservatively represented by Sales, General, Administrative and Research & Development expenses (SG&A plus R&D), as reported according to generally accepted accounting principles. In the case of banking enterprises, all expenses other than the costs of interest are, in fact, information resource costs, because a bank is essentially an information handling enterprise. Accordingly, we classify “interest expense” as the equivalent of the industrial firm’s accounting definition as “cost of goods sold” and “non-interest plus other operating expense” as equivalent of SG&A plus R&D.

4. The model and analysis for panel data

Our data set contains both cross-sectional and time series data ranging from the time period 1989 to 1997. Such data sets are called panel or longitudinal data. A simple OLS regression suffers from inefficiency, multicollinearity, and correlation between the explanatory variables and the error terms1 with the estimation being biased [14].

In Fig. 1, the oval represents observation points for a company over time. The dotted line is the regression of variable I on Y for each company. The straight line is the OLS regression. The data of each company shows that variable I has little relation with Y. It is also evident that the OLS regression overestimates its relation. It is the companies’ differences (maybe the size of the companies) represented by the different height of the three companies, related to Y. Ben-Porath gave another example [3]. Suppose that a cross-sectional sample of married women is found to have an average yearly labor force participation rate of 50%; at one extreme this might be interpreted as “each woman has a 50% chance of being in the labor force in any given year,” while at the other extreme it might

---

1 Multicollinearity is the relation between two explanatory variables. In mathematical term, it means $|x^T x| 
eq 0$. Panel data contain time series data and multicollinearity becomes possible. Efficiency is measured by the p-value. Increasing the degrees of freedom increases the efficiency and panel data has more observations than pure time series or pure cross-section data and therefore can increase the degrees of freedom. OLS also assumes the explanatory variables do not correlate with the error terms. When the relation occurs, the estimation is biased. Panel data analysis provides means to eliminate or reduce such a bias.
imply that “50% of the women always work and 50% never work” (Fig. 2).

From this, it is clear that we need to isolate the company’s specific effect to analyze panel data. To this end, our model in Cobb–Douglas production function becomes:

\[
\ln Y_{it} = \ln a_0 + \alpha_I \ln I_{it} + \alpha_N \ln N_{it} + \alpha_{IN} \ln IN_{it} + \alpha_L \ln L_{it} + \alpha_{OE} \ln OE_{it} + \mu_i + \nu_{it},
\]

A traditional Cobb–Douglas function is

\[
\ln Y_{it} = \ln a_0 + \alpha_I \ln I_{it} + \alpha_N \ln N_{it} + \alpha_{IN} \ln IN_{it} + \alpha_L \ln L_{it} + \alpha_{OE} \ln OE_{it} + v_{it},
\]

Here, the subscript \( i \) denotes the banks, and \( t \) the time. \( \nu_{it} \) is the disturbance. \( I, N, IN, L, \) and \( OE \) represent IT budget, non-interest expense, interest expense, staff (labor) cost, and operating expense. The output \( Y \) is defined as the net sales or revenue in the income statement and, therefore, the parameters can give us the estimation of the impact from each input variable on the sales. We do not use profit because it is the net of revenue from all investments. The relationship between a particular input and the profit therefore contains impact from other inputs. As such, it is difficult to explain the meaning of the parameters. However, we will see the marginal product, \( MP_i \), derived from the parameters can still give us the return on investment from input \( i \).

The disturbance in Eq. (1) is separated into two kinds. \( \nu_{it} \) is the random error. It shows the disturbance due to the unobservable difference between banks, for example, company size. This model therefore, acknowledges and controls the heterogeneity of banks.

There are several methodologies to analyze panel data. The ‘within’ model assumes that there are common slopes but that each cross section unit (i.e., each bank) has its own intercept (e.g., company size varies.) The ‘between’ model regresses only the independent variable-specific difference (e.g., bank size) on the dependent variable. The random effects model assumes that the intercepts are drawn from a common distribution with mean \( \mu \) and variance. The estimator in the ‘random effects’ model is computed by estimating the relative importance of between and within variation of the disturbance \( \mu_i - \nu_{it} \) and using this estimated ratio to combine the within and between estimators optimally [11]. We use the random effect model, to measure the output elasticity and marginal product of input variables of a Cobb–Douglas function.

The Cobb–Douglas production function satisfies some economic rules. One important rule is the law of diminishing marginal productivity. It means, for example, that the first 100 units of \( I \) can produce, say 80 units of \( Y \), but the second 100 units of \( I \) can produce only 70 units of \( Y \). It is clear that a Cobb–Douglas function can satisfy this property when \( \alpha_i < 1 \). Because a production function has to satisfy this basic economic property, we cannot simply use a linear regression. This Cobb–Douglas function also satisfies some other economic properties; for example, the law of diminishing marginal rate of technical substitution. However, it is not the purpose of this paper to prove that it can satisfy most economic properties. We are concerned with the use of an appropriate analysis for panel data to the extent that the results will be statistically unbiased and consistent.

From Eq. (1), we can calculate the output elasticity. This is used to measure the percentage change of the
output due to one percentage change of an input. Therefore, it is defined as $\ln Y / \ln X_i$, or $\frac{\partial Y}{\partial X_i}$. We also know that the parameters $\alpha_I$, $\alpha_N$, $\alpha_{IN}$, $\alpha_L$, and $\alpha_{OE}$ in this logarithmic transformed Cobb–Douglas function are actually the measurements for $\ln Y / \ln X_i$, and are therefore the output elasticity for each input variable. From the output elasticity, we can calculate the marginal product from this equation as: $MP_i = \alpha_i Y / X_i$, since the marginal product is $\partial Y / \partial X$. The marginal product is important because it measures how many units of an output will be increased by increasing one unit of an input. In our case, if $MP_i$ is greater than one, it is clear that input $i$ on the margin generates positive profit, and if it is less than one, an additional dollar of input $i$ does not provide more value than US$ 1.

5. Finding

First, let us compare the estimation between OLS and random effect model for panel data Table 1. Panel data analysis shows positive IT productivity illustrates the results from random effect model while Table 2 is from OLS.

Although some estimators are significant in Table 2, the estimation for IT and labor are not satisfactory. Neither is significant under a 1% confidence interval. The $p$-value for labor from OLS is extremely large—0.784, compared with 0.159 from the random effect model. Therefore, a random effect model is super both in theory and in empirical results when using panel data.

We can calculate the marginal product for each parameter by the formula $MP_i = \alpha_i Y / X_i$, where $\alpha_i$ is the output elasticity of variable $i$. The result is found in Table 3.

It shows that information spending has the highest marginal product among all the input variables. Among them, labor stands as the lowest productive input. Since these two variables are mostly substitutable, the difference between their marginal products may justify a banks' decisions to substitute labor for IT products. Since MP means how much the dollar value increases in output for an additional US$ 1 input, we can see that IT is the only variable that provides more than a dollar return for a dollar expense on the margin. This would suggest increasing IT investment. We also see that for each additional dollar of labor expense, the output value can increase by only 18 cents, so we would suggest cutting labor expenses. This suggestion is strengthened by the fact that the average spending on IT is far less than other inputs (see column Average of Table 3).

This shows a sharp difference from previous IT productivity reports based on the data of the 1980s. Even though many studies have shown strong IT productivity, the paradox cannot be cleared without a positive result in the service industry. The 1980 data may not reveal strong IT productivity in the banking industry but this research, using recent data, shows that IT is the only variable with positive marginal gain and its productivity is far better than labor.

6. Conclusion

The purpose of this paper was to try to enhance our understanding of the IT productivity paradox in the banking industry. This issue became more important...
in the 1990s when journalists and economists alike observed a consequence of the so-called ‘New Economy’ [22]. Articles from Business Week have provided evidence that the nature of our economy now is different from before and one of the fundamental drives is IT. Shepard stated, “... information technology accounts for a quarter to a third of economic growth...” Statistics have shown that productivity was weak from the 1960s to the early 1990s, but there are signs of a revival; e.g., the U.S. Department of Labor data showed that the average productivity growth rate dropped from 2.68% in the 1960s to 1.26% in 1980s but it bounced back to 1.46% in the 1990s up to 1998.

Yet, there has been no evidence showing that the recent productivity growth reversal is due to IT. The only way to prove that it is playing a significant role in improving productivity is to provide an accurate methodology of measuring IT productivity. Here, we have used bank data from 1989 to 1997. The data covers section sectional and time series data, so we have applied the random effect model—the model control individual (bank) heterogeneity. IT productivity, precisely speaking, IT contributions to corporate profit was measured. Our production function shows that IT is the only input variable that provides more dollar value than the input cost on the margin when it is compared with interest expense, non-interest expense, staff cost, and operating expense. For those New Economy advocates and those who claimed the IT productivity paradox has been solved, this paper is encouraging, because it suggests that IT maybe one of the positive drivers for recent productivity gains by large U.S. companies.

One must caution against making generalized conclusions about productivity from a limited sample of banking data. Financial institutions are unique in their extraordinary dependency on IT spending as compared with staff compensation costs.

The median compensation of banking employees in our sample is US$ 51,500 which is well above the median for industrial firms. The information-intensive banking operations are also more amenable to routine standardization to an extent that is rarely found in other sectors of the economy. For these reasons, our favorable findings about the benefits of substituting expensive labor costs with IT are not surprising.

The contribution of this paper is not only its discovery achieved by using more recent data, but also the methodology proposed. The data set can be biased if it contains both time series and cross-section data and when an OLS regression is applied. We have carefully chosen the appropriate methodology so that the results can be more accurate and convincing. One important aspect is the endogeneity of the input variables. This argues that the input variables are not independent. A firm’s objective is to maximize profit or minimize cost and it will adjust the output and input volumes to achieve this goal. The change of the inputs arises in large part because of intentional actions taken by firms who respond to market incentives and as Varian and Intriligator suggested, a model incorporating this assumption is better than one without [15,31] both in an economics and in econometric sense. One important criterion for firms to adjust the input quantity is price. When IT prices drop and other input prices are unchanged, companies would buy more IT and reduce their buying on other input factors.

References