RIVISTA INTERNAZIONALE DI SCIENZE ECONOMICHE E COMMERCIALI

RISEC

INTERNATIONAL REVIEW OF ECONOMICS AND BUSINESS

Vol. L - No. 4 – December 2003

Pubblicazione trimestrale - Sped. in a.p. - 45% - art. 2 comma 20/B - legge 662/96 - Filiale di Bologna

Profitable Technical Trading Rules for the Italian Stock Market
MASSOUD METGHALCHI and YUNG-HO CHANG 433

The Demand for Money by Private Firms in a Regulated Economy: Theoretical Underpinnings and Empirical Evidence for Germany, 1960-1998
MARTIN T. BOHL and FRIEDRICH L. SELL 451

Quality of Assets, Bank Profitability and Capital Position
ATHANASIOS G. NOULAS and JOHN A. PAPANASTASIOU 475

Operational Efficiency in the Greek Banking Industry: Evidence from the Stochastic Frontier Approach and Error Decompositions
NICHOLAS APERGIS and ANTHONY REZITIS 485

Stock Price and Exchange Rate Interactions in an Emerging Market
WAN N.W. AZMAN-SAINI, MAZUFAR S. HABIBULLAH, and MOHAMED AZALI 503

The Early Development of Banking in Ethiopia ARNALDO MAURI 521

Management Quality, Credit Policy, and Loan-Loss Differences among Commercial Banks in Nigeria
RALPH I. UDEGBUNAM 545

Foreign Direct Investment and Economic Growth in Sub-Saharan Africa
ANTHONY E. AKINLO 569

The Impact of Guyana's Foreign Debt Burden on Its Rate of Economic Growth, 1971-2001
NELSON C. MODESTE and HECTOR BUTTS 581

UNDER THE AUSPICES OF BOCCONI UNIVERSITY AND THE UNIVERSITY OF MILAN

CEDAM

RIVISTA INTERNAZIONALE DI SCIENZE ECONOMICHE E COMMERCIALI

RISEC

INTERNATIONAL REVIEW OF ECONOMICS AND BUSINESS

(a quarterly journal founded by Giovanni Demaria with Tullio Bagiotti)

EDITOR

ALDO MONTESANO (Bocconi University, Milan)

EDITORIAL BOARD

ACHILLE AGNATI (University of Padua) - PIERPAOLO BATTIGALLI (Bocconi University, Milan) - GIOVANNI BELLONE (University of Padua) - GIUSEPPE BOGNETTI (University of Milan) - DAVIDE CANTARELLI (University of Padua) - MICHEL DE VROEY (Catholic University of Louvain) - FRANCO DONZELLI (University of Milano-Bicocca) - MICHELE GRILLO (Catholic University, Milan) - ANTON MUSCATELLI (University of Glasgow) - PIER LUIGI PORTA (University of Milano-Bicocca) - DOMENICO SARTORE (University of Venice) - LUCA STANCA (University of Milano-Bicocca) - PIERO TEDESCHI (University of Milano-Bicocca) - EDWARD TOWER (Duke University)

Managing Editor: ANNA BAGIOTTI CRAVERI
Editorial Office: Via P. Teuliè 1, 20136 Milano (Italy) - Tel. & Fax: +39-02-58317434 - E-mail: risec@tin.it

ADVISORY BOARD

MAURICE ALLAIS (Ecole Nationale Supérieure des Mines, Paris) - MARIO ARCELLI (Luiss, Rome) - HENRI BARTOLI (University of Paris) - WILLIAM J. BAUMOL (Princeton University) - MARTIN J. BECKMANN (Technische Universität München) - MARK BLAUG (University of Exeter) - FINN FORSUND (Oslo University) - WILLIAM D. GRAMPP (University of Chicago) - LENNART HJALMARSSON (Göteborg University) - ARNALDO MAURI (University of Milan) - HISAO ONOE (Kobe University) - ALBERTO QUADRIO CURZIO (Catholic University, Milan) - ROBERTO RUOZI (Bocconi University, Milan) - ROBERT M. SOLOW (M.I.T., Cambridge, MA) - SERGIO STEVE (University of Rome) - ROBERT SUGDEN (University of East Anglia) - MARIO TALAMONA (University of Milan) - BASIL YAMEY (London School of Economics) - STEFANO ZAMAGNI (University of Bologna)

CISQCERT UNI EN ISO 9001

La Casa Editrice CEDAM S.p.A. opera con un Sistema Qualità conforme alle norme UNI EN ISO 9001 certificato da CISQCERT con numero 1.354

Direttore responsabile: Aldo Montesano - Autorizz. Tribunale di Treviso N. 113 del 22-10-54

USPI Rivista associata all'Unione della Stampa Periodica Italiana

Leonelli Printing s.r.l. - Villanova di Castenaso (Bo)

Copyright - Printed in Italy

RISEC, Volume 50 (2003), No. 4, 433-450

PROFITABLE TECHNICAL TRADING RULES FOR THE ITALIAN STOCK MARKET

by
MASSOUD METGHALCHI* and YUNG-HO CHANG**

1. *Introduction*

Fama (1970) defined an efficient financial market as one in which security prices always fully reflect the available information; any new information will be quickly and instantaneously reflected in prices. Furthermore, since news on any company, by definition, is unpredictable (arrives randomly), price changes will be unpredictable or follow a random walk. The "weak-form" Efficient Market Hypothesis (EMH) asserts that stock prices already reflect all information that can be derived by investigating market trading data such as the history of past prices or trading volume. Advocates of the weak-form market efficiency hypothesized that investors could not drive profits above a buy-and-hold strategy using any trading rule that depended solely on past market information such as price or volume, implying that technical trading rules are useless.

After more than three decades of research and literally thousands of journal articles, financial economists and practitioners have not yet reached a consensus whether technical trading rules could discern recurring-price patterns for profitable trading. The overwhelming majority of financial economists support the "weak-form" efficient market hypothesis. This is because much of earlier research supported the random walk hypothesis. The

* School of Business, University of Houston-Victoria, Texas, (USA). E-mail: metghalchim@uhv.edu

** Department of Finance, Tunghai University, Taiwan.

The authors extend their thanks to the referees for their invaluable and constructive suggestions.

following studies have long supported the weak-form market efficiency: Larson (1960), Osborne (1962), Alexander (1964), Granger and Morgenstern (1963), Mandelbrot (1963), Fama and Blume (1966). Fama (1965) finds that stock prices indeed follow random walks and he finds no systematic evidence of profitability of technical trading strategies. Van Horn and Parker (1967) using a simple technical trading rule for 30 NYSE securities confirm the random walk hypothesis. Jensen and Benington (1970) conclude that technical trading rules are not useful. By the end of the 1970s, the attitude of academics towards technical trading is described by Shleifer (2000, p. 9): "as matters stood at the end of the 1970s, the EMH was indeed one of the great triumphs of twentieth-century economics."

However, since the early1990s, technical trading has been enjoying a renaissance both on Wall Street and in academic circles. Several papers have presented evidence that some simple trading rules are useful for predicting stock market returns. The cornerstone of this new research on technical analysis is an article by Brock, Lankonishok and LeBaron (1992); BLL analyzed moving averages and trading range breaks on the Dow Jones Industrial Index from 1897 to 1985. They used various short and long moving averages of prices to generate buy and sell signals. They tested long moving averages of 50, 150 and 200 days with short averages of 1, 2 and 5 days. They point out that "all buy-sell differences are positive and the *t*-tests for these differences are highly significant..." and they go on to conclude that their "results are consistent with technical rules having predictive power". Other researchers have used some variants of BLL's moving averages to investigate whether stock market indices can be predicted by some simple form of technical analysis. Bessembinder and Chan (1995) conclude that the BBL's rules are successful in predicting stock price movements in Japan, Hong Kong, South Korea, Malaysia, Thailand and Taiwan, with the predictability strongest in the last three markets. Ergul, Holmes and Priestley (1997), using daily closing prices of 63 stocks traded on the Istanbul Stock Exchange, conclude that technical analysis on volume can aid the prediction of returns which cannot be predicted by the analysis of past returns in isolation. Pruitt and White (1998), using the University of Chicago's CRSP daily data tapes over the 1976-1985 period, conclude that technical trading rules are capable of outperforming a simple buy-and-hold strategy even adjusting for transaction costs. Bessembinder and Chan (1998) confirm the basic BLL results; however, they argue that the BLL results can coexist with the notion of market efficiency when considering transaction costs. Gencay (1998a, 1998b) and Ratner and Leal (1999) also support the predictive power of technical trading rules. Kwon and Kish (2002), applying three popular

technical trading rules to NYSE index over the period 1962-1996, conclude that the technical trading rules have the potential to capture profit opportunities over various models when compared to the buy-and-hold strategy. However, in a recent study Ready (2002) points out that the apparent success of the BLL moving average rules is a spurious result of data snooping and need not persist in the future. Technical trading rules have also been applied to foreign exchange markets. For a survey of technical trading on foreign exchange markets see Taylor and Allen (1992) and Maillet and Michel (2000).

In this paper, we investigate whether some form of technical analysis can predict changes in the Italian Stock Market (ISM) index. Secondly, if the Italian stock market exhibits substantial deviations from random walk hypothesis, can we identify a trading strategy that will beat the buy-and-hold strategy? Although many earlier studies concluded that trading rules are useless, the recent studies on technical analysis suggest that this conclusion might have been premature. We provide empirical evidence by evaluating many technical trading rules to shed additional light on this controversy. The ISM is interesting for two reasons. First, because of inflation and regulations in the 1970s and 1980s, the ISM was small compared to GDP (market capitalization of the ISM as a percent of GDP were 5.0%, 13.1%, 18.6%, and 70.2% at the end of 1975, 1985, 1995, and 2000 respectively according to Borsa Italiana); however, since the 1990s, because of extensive privatization and deregulation carried out by Italy's government, the ISM has grown substantially and at the end of 2000 the ISM was worth 70.2% of Italy's GDP, in line with capitalization/GDP ratio of Germany but still lower than France and UK. During the 1990s, far-reaching measures have improved the organization of the market, considerably increasing the liquidity, efficiency, and transparency of the ISM. Second, since the ISM, regarding major economic variables and technology of the financial market, is comparable with those of the other European financial markets, we think the result of our research could be of interest for other European financial markets. This paper is structured as follows: Section 2 discusses data and methodology, Section 3 presents empirical results on technical trading rules, Section 4 compares various strategies with the buy-and-hold strategy, and finally Section 5 concludes on this controversy.

2. *Data and Methodology*

We use Datastream's daily closing price index of the Italian stock market

over the period of January 2, 1973 to December 31, 2002 and compute daily returns as changes in logarithms of the stock index level. Although changes in stock price index do not include daily dividend yields, we do not expect this omission to alter the results of our analysis. We estimate approximate annualized returns on the basis of 260 trading days per year as exponential (260R) −1, where R is the average daily return. In Section 4, when we discuss trading strategies, daily money market rates are needed to estimate returns of various strategies. The monthly data on money market rates for Italy are taken from IFS of International Monetary Fund. In order to get the approximate daily overnight rates, we divide monthly yield by 365. (We could have divided by 260, or 52 weeks with 5 working days each; however, that would overestimate the returns of various strategies that we will discuss in Section 4).

Technical analysis is based on the idea that prices move in trends, which are determined by the changing attitudes of traders towards various economic, political and psychological forces. As Pring (1991, p. 2) points out, "the art of technical analysis is to identify trend changes at an early stage and to maintain an investment posture until the weight of evidence indicates that the trend has reversed". One of the most important Trend-Determining Techniques is based on the crossing of two moving averages (*MA*) of prices. According to this rule, buy (sell) signals are emitted when the short-term moving average exceeds (is less than) the long-term moving average by a specified percentage. In this study we use long moving averages of 50, 100, 150 and 200. As for the short moving average, like the BLL study, we use 1 day (the raw price) moving average. (We have also used 5 and 10 days short *MA*; the results were not significantly different from 1 day). Thus, a buy signal is emitted when the short moving average breaks the long one from below and a sell signal is emitted when the short moving average breaks the long from above.

We define P_t as the short moving average or the raw index level at time t, and define long moving average of M at time t as:

$$MA_t(M) = \frac{1}{M} \sum_{i=0}^{M-1} P_{t-i} \tag{1}$$

We will be either in the market (buy days) or out of the market (sell days). We assume that a trader following these *MA* strategies could presumably observe the prices a few minutes prior to the day's close and make the trading decision at the close. If the closing price is above the long moving average (or above 1.01, 1.02 times long-run moving average), then the trader will be

in the market next day by buying the index at the closing price (next day will be a buy day). Next day's return will be the difference between the logarithm of the closing price next day and the logarithm of closing price the previous day. On the other hand, if the closing price is below the long moving average (or below 1.01, 1.02 long *MA*), then we will sell the index at the closing price and will be out of the market next day (sell days). We define mean buy and mean sell returns as follows:

$$X(b) = \frac{1}{N_b} \sum R_b \tag{2}$$

$$X(s) = \frac{1}{N_s} \sum R_s \tag{3}$$

where N_b and N_s are the total number of buy and sell days and R_b and R_s are daily returns of buy and sell days. We will test whether the returns of any moving average trading rules are greater than a buy-and-hold strategy and whether the mean buy is different than the mean sell. The null and alternative hypotheses are expressed as:

$H_0 : X(b) - X(h) = 0,\ X(s) - X(h) = 0\ ,\ X(b) - X(s) = 0$
$H_A: X(b) - X(h) \neq 0,\ X(s) - X(h) \neq 0,\ X(b) - X(s) \neq 0$

Where $X(h)$ is the mean return for the buy-and-hold strategy. The test statistic for the mean buy returns over the mean buy-and-hold strategy is:

$$t = \frac{X(b) - X(h)}{\sqrt{Var(b)/N_b + Var(h)/N}} \tag{4}$$

where $Var(b)$ and $Var(h)$ are the variance of buy and buy-and-hold returns respectively. The above formula is also used to test the mean sell returns over the mean buy-and-hold strategy; and the mean buy returns over the mean sell returns by replacing the appropriate variables in the *t*-statistic formula.

3. *Empirical Results*

In Table 1 we present the summary statistics for the buy-and-hold strategy for the entire 30 years and for three 10-year sub-periods. As we can see from Table 1, for the entire period the daily average of buy-and-hold strategy is .036 percent per day with a standard deviation of .01361. The *t*-value for the

entire period is equal to 2.33 (.00036 divided by .01361/$\sqrt{7776}$). The annual average is 9.81 % per annum for the full sample. In this paper, we compare all *t*-stats with 1.96, the critical *t*-value at 5 percent level for large numbers of observations. The unconditional mean for the entire period is significantly different from zero; however, we cannot reject the hypothesis that the unconditional means for the three sub-periods are zero.

Table 1. Summary statistics for daily returns

	Daily mean	*SD*	*N*	*t*-stat	Annual average
3/13/73 to 12/31/02	.00036	.01361	7776	2.33*	9.81 %
3/13/73 to 12/31/82	.00034	.01475	2559	1.17	9.24 %
1/3/83 to 12/31/92	.00043	.01215	2609	1.81	11.83 %
1/4/93 to 12/31/02	.00030	.01384	2608	1.11	8.11 %

Results are presented for the full sample and 3 non-overlapping sub-periods. Returns are measured as log differences of the level of stock index, *t*-stat tests whether mean daily returns are different from zero. Numbers marked with asterisk are significant at the 5 % level for a two-tailed test. Annual mean is estimated by exp(260*daily return) – 1. *SD* stands for standard deviation and *N* represents number of observations.

Table 2 summarizes the results of standard moving average trading rules. Most of these trading rules have been used in the BLL study. The rules are described as (1, long, percentage). For each rule we report mean returns on buy days and sell days, standard deviations of returns on buy and sell days, and total number of buy and sell days. The numbers in the parentheses are the *t*-statistics (equation 4) testing the difference of the mean buy and mean sell from the unconditional 1-day mean, and buy-sell from zero.

The first row of Panel A reports results of trading rule (1,50,0); we will be in the market (buy days) if the *MA*1 (price level) is greater than *MA*50 and out of the market (sell days) if *MA*1 is less than or equal to *MA*50. The last row of Panel A shows the results of trading rule (1,200,2); we will be in the market (buy days) if *MA*1 is greater than 1.02**MA*200, otherwise out of the market (sell days). The results of Table 2 are amazing. All the buy-sell differences (column 4) are positive and the *t*-stats for these differences are highly significant, rejecting the null hypothesis of equality with zero. The mean buy and sell returns are shown in columns 2 and 3. The buy returns are all positive with an average one-day return of .00093.This compares with the unconditional 1-day return of .00036 of Table 1. Five of the twelve tests reject the null hypothesis that the buy returns equal the unconditional 1-day return.

Table 2. Standard moving average rules

Rules	Buy	Sell	Buy–Sell	SD_b	SD_s	N_b	N_s
Panel A: 1973–2002							
(1,50,0)	0.00104 (2.75)*	–.00045 (–2.8)*	0.00149 (4.78)*	0.01274	0.01453	4209	3567
(1,50,1)	0.00123 (3.29)*	–.00075 (–2.89)*	0.00197 (5.35)*	0.01301	0.01409	3663	4113
(1,50,2)	0.00139 (3.61)*	–.00034 (–2.76)*	0.00173 (5.52)*	0.01361	0.01366	3134	4642
(1,100,0)	0.00078 (1.75)	–.00022 (–1.91)	0.00100 (3.15)*	0.01301	0.01438	4321	3405
(1,100,1)	0.00087 (2.05)*	–.00022 (–2.02)*	0.00109 (3.52)*	0.01318	0.01408	3943	3783
(1,100,2)	0.00097 (2.32)*	–.00020 (–2.03)*	0.00117 (3.77)*	0.01330	0.01389	3568	4158
(1,150,0)	0.00079 (1.74)	–.00022 (–1.98)*	0.00101 (3.20)*	0.01296	0.01431	4388	3288
(1,150,1)	0.00084 (1.87)	–.00019 (–1.95)	0.00103 (3.30)*	0.01314	0.01401	4094	3582
(1,150,2)	0.00083 (1.77)	–.00010 (–1.69)	0.00093 (3.00)*	0.01334	0.01376	3788	3888
(1,200,0)	0.00077 (1.64)	–.00020 (–1.91)	0.00097 (3.06)*	0.01323	0.01395	4344	3282
(1,200,1)	0.00082 (1.81)	–.00020 (–1.95)	0.00101 (3.24)*	0.01329	0.01384	4104	3522
(1,200,2)	0.00085 (1.90)	–.00018 (–1.94)	0.00103 (3.32)*	0.01339	0.01371	3915	3711
Average	0.00093 (2.21)*	–.00027 (–2.15)*	0.00120 (3.77)*	0.01318	0.01402	3956	3745
Panel B: sub–periods							
73–82: (1,50,0)	.00131 (2.03)*	–.00080 (–2.13)*	.00211 (3.60)*	.01414	.01536	1378	1181
83–92: (1,50,0)	.00133 (2.33)*	–.00058 (–2.31)*	.00191 (4.00)*	.01130	.01297	1382	1227
93–02: (1,50,0)	.00051 (.49)	.00004 (–.50)	.00047 (.84)	.01263	.01522	1449	1159

Results for daily data from 1973-2002. Rules are identified as (short, long, band) where short and long are the short and long moving averages and band is percentage difference to generate a signal. N_b and N_s are the numbers of buy and sell signals reported in each period SD_b and SD_s are standard deviations of buy and sell signals, respectively. The numbers in the parentheses are the t-statistics testing the difference of the mean buy and mean sell from the unconditional 1-day mean, and buy-sell from zero. Numbers marked with asterisks are significant at the 5% level for a two-tailed test.

For the sells, the results are even stronger. All the sell returns are negative with a mean daily return of –.00027. Six of the twelve tests reject equality with the unconditional mean return and the other six are marginally significant. The standard deviations of buy days and sell days are reported in Columns 5 and 6. The standard deviations for buy days are always smaller than those for sell days for the full sample. This implies that the market is less volatile for buy periods than for sell periods. Columns 7 and 8 report the number of buys and sells for various rules. On the average, 51.4 % of the time we are in the market (buy days) and 48.6 % out of the market (sell days). As can be seen from Table 2, our best results (high *t*-stats) are with moving average 50s. We also tested all 12 rules by replacing short moving average of one day with short moving average of 5 and 10 days, a total of 24 tests, the results were similar to *MA*1.

The negative returns in Table 2 for sell days are especially noteworthy. As BLL point out, these returns cannot be explained by various seasonalities since they are based on 48.6 % of all trading days. This predictability of returns can reflect either (1) changes in expected returns generated from an equilibrium model, or (2) market inefficiency. Although changes in expected returns are possible, it is hard to imagine an equilibrium model that predicts negative returns over such a large fraction of trading days.

In Panel B of Table 2 we present the results for three sub-periods. To save space, results are presented for only the (1,50,0) rule. The results are mixed over different sub-periods. While the *t*-values of buy, sell, and buy-sell are significant at 5% level over the first two sub-periods, the *t*-values are not statistically significant for the period 1993-2002. This implies that the standard moving average trading rules does not outperform the buy-and-hold strategy in the third period of this study, therefore, we could conclude that the Italian stock market has become more efficient over the past decade.

Given that the standard trading rules of BLL did not work for the last decade for the Italian stock market, we consider the moving average technical trading rules advocated by Arnold and Rahfeldt, (AR), (1986). As they point out (p. 71), the rules might then be: "Buy when the actual price crosses above both moving averages and exit the market when the price crosses below either market moving average". The AR trading rules compare the price level with two moving averages: moving averages short and long. If the price index is above both moving average short and moving average long, the rule emits a buy signal and we will be in the market (buy days); we will be out of the market (sell days) if the price is below either moving average. Therefore the AR rules can be summarized as follows:

If $P > MA(S)$ and $M(L)$ we will be in the market (buy days)
If $P <$ either $MA(S)$ or $MA(L)$, we will be out of the market (sell days)

where $MA(S)$ and $MA(L)$ represent moving average short and moving average long respectively. We will consider the AR trading rules for short moving average of 5 and 10 days and long moving average of 50,100,150 and 200 days.

Table 3 presents the results of AR moving average trading rules. The first row of Panel A in Table 3 reports results of trading rule of $P >$ (5,50); we will be in the market (buy days) if the price level is greater than both short moving average of 5 days and long moving average of 50 days and out of the market (sell days) if the price level is less than either moving average. The second row shows the result of AR trading rule for short moving average of 5 days and long moving average of 100 days. We will be in the market (buy days) if price is greater than both moving averages of 5 and 100 days and out of the market (sell days) if price is less than either moving average 5 or 100 days. The results of Table 3 are spectacular and are much stronger than the results of Table 2. Panel A presents results for the full sample. The t-values for all buy and all sell are highly significant, rejecting the null hypothesis that the buy and sell returns equal the unconditional 1-day return. All the buy-sell differences are positive and the t-stats for these differences are highly significant, rejecting the null hypothesis of equality with zero.

In Panel B of Table 3 we report the results for three sub-periods for AR trading rules. For all sub-periods, the test statistics reject the null hypothesis that the buy returns equal the unconditional buy-and-hold returns. All the buy-sell differences are positive and the t-stats for these differences are highly significant, rejecting the null hypothesis of equality with zero. For AR trading rules, it is noteworthy to observe that the number of sell days is higher than the number of buy days; on the average, 36% of the time we are in the market (buy days) and 64% out of the market (sell days).

If technical analysis did not have any power to forecast price movements, then we should observe that the buy days returns do not differ appreciably from sell days returns. However, the results of Table 2 and 3 indicate that moving average rules do indeed have predictive power and can discern recurring-price patterns for profitable trading. Given this predictive power of technical analysis, can we design various trading strategies to beat the buy-and-hold strategy?

Table 3. AR moving average rules

Panel A: 1973-2002							
Rules	Buy	Sell	Buy–sell	SD_b	SD_s	N_b	N_s
$P > (5.50)$	0.00209 (6.14)*	–.00059 (–3.76)*	0.00268 (8.67)*	0.01236	0.01417	2751	5025
$P > (5.100)$	0.00196 (5.67)*	–.00051 (–3.36)*	0.00246 (7.91)*	0.01234	0.01420	2658	5068
$P > (5.150)$	0.00189 (5.37)*	–.00044 (–3.17)*	0.0023 (7.48)*	0.01232	0.01410	2621	5055
$P > (5.200)$	0.00188 (5.27)*	–.00042 (–3.08)*	0.00230 (7.31)*	0.01241	0.01403	2556	5070
$P > (10.50)$	0.00172 (4.99)*	–.00050 (–3.33)*	0.00222 (7.27)*	0.01237	0.01428	3008	4768
$P > (10.100)$	0.00196 (5.67)*	–.00051 (–3.36)*	0.00247 (7.91)*	0.01234	0.01419	2658	5068
$P > (10.150)$	0.00154 (4.25)*	–.00031 (–2.59)*	0.00185 (5.99)*	0.01218	0.01423	2749	4927
$P > (10.200)$	0.00160 (4.20)*	–.00032 (–2.65)*	0.00193 (6.19)*	0.01225	0.01416	2662	4964
Average	0.00183 (5.20)*	–.00045 (–3.17)*	0.00228 (7.34)*	0.01232	0.01417	2708	4993
Panel B: sub-periods							
73–82: Average	.00187 (2.74)*	–.00046 (–1.64)	.00234 (3.83)*	.01394	.01504	818	1666
83–92: Average	.00207 (3.82)*	–.00050 (–2.36)*	.00257 (5.41)*	.01106	.01263	951	1659
93–02: Average	.00145 (2.40)*	.00004 (–.1.42)	.00047 (3.35)*	.01204	.01472	939	1670

Results for daily data from 1973–2002. Rules are identified as buy whenever $P >$ (short, long). For example, rule $P >$ (5,50) means: buy if the price level is greater than both short moving average of 5 days and long moving average of 50 days and sell if the price level is less than either moving average. N_b and N_s are the numbers of buy and sell signals reported in each period. SD_b and SD_s are standard deviations of buy and sell signals, respectively. The numbers in the parentheses are the t-statistics testing the difference of the mean buy and mean sell from the unconditional 1-day mean, and buy-sell from zero. Numbers marked with asterisks are significant at the 5% level for a two-tailed test.

4. *Trading Strategies*

We next provide some information on the degree to which traders using these technical trading rules can earn trading profits that could beat the buy-and-hold strategy. Given that the mean buy is greater than the mean sell and the unconditional 1-day mean, the profitability of technical trading rules

depends on trading strategy, especially, what position should the trader take when the rule emits sell signals. If the trader does not invest on the sell days, then the trader's return on the sell days will be zero which will result in a mean return of $(N_b/N)*X(b) + (N_s/N) *0$ for this strategy. In this study, we consider three strategies: (1) the trader will be in the stock market when trading rules emit buy signals and be in the money market when it emits sell signals (long/money); (2) the trader will be in the stock market when trading rules emit buy signals and short the stock market when it emits sell signals (long/short); and (3) the trader will borrow at the money market rate and double stock investment when trading rules emit buy signals and be in the money market when it emits sell signals (leverage/money). The total trading return on buy days of this strategy is $TR_t = 2 * R_t - M_t$, where R_t is the index return on day t and M_t is the daily money market rate.

For each strategy we estimate the daily return and then subtract from it the daily return from the buy-and-hold strategy to get the daily difference return. To test whether the average daily difference is greater than zero, we express the null and alternative hypotheses as follows:

$H_0 : \ ddif = 0$
$H_A : \ ddif > 0$

The t-statistic for the above test is:

$$t = \frac{X(ddif)}{\sqrt{Var(ddif)/N}} \tag{5}$$

where $X(ddif)$ is the average daily difference of returns of each strategy over the buy-and-hold strategy and $Var(ddif)$ is the variance of daily difference returns, and N is the total number of days. Table 4 and 5 report the results of various strategies for standard and AR moving average technical trading rules respectively.

As can be seen from Table 4, two strategies (1 & 3) beat the buy-and-hold strategy for the full sample. We cannot reject, however, the hypothesis that the mean daily difference for strategy 2 is different from zero. The standard deviation of the first strategy (.00948) is lower than the standard deviation of the buy-and-hold strategy (.01361), and the standard deviation of the third strategy (.01358) is similar to the buy-and-hold strategy. Therefore, strategies 1 & 3 have higher average returns but not higher risks, they beat the market for the full sample. However, for the three sub-periods the results are not as strong as for the full sample. Strategies 1 & 3 both beat the market for the sub-periods of 73-82 and 83-92. But for the sub-period 92-02, no strategies

beat the market. It seems that the Italian stock market has become more efficient over the past decade. The second strategy (long/short) does not perform as well as other strategies because of its restrictive assumption. We are assuming that when out of the market and shorting the market we are not earning any interest income.

Table 4. Trading strategies for standard moving average

Panel A: 1973-2002							
	Strategy 1 Long/money		Strategy 2 Long/short		Strategy 3 Leverage/money		
Rules	*X*(*ddif*)	*SD*(*ddif*)	*X*(*ddif*)	*SD*(*ddif*)	*X*(*ddif*)	*SD*(*ddif*)	*N*
(1.50.0)	0.00053 (3.16)*	0.00985	0.00042 (1.86)	0.01969	0.00075 (4.85)*	0.01359	7776
(1.100.0)	0.00024 (2.17)*	0.00955	0.00019 (0.88)	0.01909	0.00049 (3.19)*	0.01362	7726
(1.150.0)	0.00023 (2.17)*	0.00937	0.00019 (0.90)	0.01873	0.00050 (3.24)*	0.01355	7676
(1.200.0)	0.00022 (2.13)*	0.00916	0.00017 (0.82)	0.01831	0.00048 (3.09)*	0.01354	7626
Average	0.00026 (2.41)*	0.00948	0.00024 (1.12)	0.01896	0.00056 (3.59)*	0.01358	
Panel B: sub-periods							
73–82: (1.50.0)	0.00056 (2.70)*	0.01045	0.00074 (1.79)	0.02088	0.00106 (3.64)*	0.01472	2559
83–92: (1.50.0)	0.00045 (2.57)*	0.00890	0.00055 (1.57)	0.01779	0.00095 (3.99)*	0.01211	2609
93–02: (1.50.0)	0.00006 (0.29)	0.01014	–0.00004 (–0.10)	0.02029	0.00024 (.90)	.01384	2608

X(*ddif*) and *SD*(*ddif*) are average and standard deviation of daily difference between the return of each strategy and the buy-and-hold strategy. The numbers in the parentheses are the *t*-statistics testing the equality of average daily difference from zero. Numbers marked with asterisks are significant at the 5% level for a one-tailed test.

In Table 5, we report the results of the same three strategies (long/money, long/short, leverage/money) using the AR moving average technical trading rules. For AR moving average rules, the results are stronger than standard moving average rules. For the full sample, all three strategies significantly beat the buy-and-hold strategy. All three *t*-values are significant at 5 % level, rejecting the null hypothesis that the mean return of each strategy is equal to the mean buy-and-hold strategy.

Table 5. Trading strategies for AR moving average rules

Panel A: 1973-2002							
	Strategy 1 Long/money		Strategy 2 Long/short		Strategy 3 Leverage/money		
Rules	$X(ddif)$	$SD(ddif)$	$X(ddif)$	$SD(ddif)$	$X(ddif)$	$SD(ddif)$	N
$P > MA(5,50)$	0.00059 (4.53)*	0.01140	0.00076 (2.95)*	0.02278	0.00121 (7.88)*	0.01356	7776
$P > MA(5,100)$	0.00054 (4.14)*	0.01150	0.00067 (2.55)*	0.02299	0.00111 (7.15)*	0.01359	7726
$P > MA(5,150)$	0.00050 (3.82)*	0.01145	0.00058 (2.22)*	0.02288	0.00104 (6.71)*	0.01352	7676
$P > MA(5,200)$	0.00049 (3.76)*	0.01145	0.00056 (2.14)*	0.02289	0.00102 (6.57)*	0.01351	7626
$P > MA(10,50)$	0.00050 (3.95)*	0.01119	0.00062 (2.43)*	0.02237	0.00104 (6.78)*	0.01357	7776
$P > MA(10,100)$	0.00044 (3.41)*	0.01141	0.00049 (1.87)	0.02282	0.00091 (5.87)*	0.01360	7726
$P > MA(10,150)$	0.00040 (3.06)*	0.0141	0.00039 (1.51)	0.02281	0.00084 (5.41)*	0.0354	7676
$P > MA(10,200)$	0.00042 (3.18)*	0.0143	0.00042 (1.60)	0.02285	0.00086 (5.58)*	0.0353	7626
Average	0.00049 (3.73)*	0.01141	0.00056 (2.16)*	0.02280	0.00100 (6.49)*	0.01355	
Panel B: sub-periods							
73-82: $P > MA(5,50)$	0.00079 (3.25)*	0.01225	0.00105 (2.17)*	0.02447	0.00152 (5.23)*	0.01468	2559
83-92: $P > MA(5,50)$	0.00065 (3.30)*	0.01009	0.00082 (2.09)*	0.02017	0.00136 (5.74)*	0.01207	2609
93-02: $P > MA(5.50)$	0.00032 (1.39)	0.01176	0.00042 (0.91)	0.02352	0.00077 (2.84)*	.01382	2608

$X(ddif)$ and $SD(ddif)$ are average and standard deviation of daily difference between the return of each strategy and the buy-and-hold strategy. The numbers in the parentheses are the t-statistics testing the average daily difference from zero. Numbers marked with asterisks are significant at the 5% level for a one-tailed test.

Advocates of the weak-form market efficiency hypothesize that investors could not drive profits above a buy-and-hold policy using any trading rule that depends only on past market information such as price. The results of Panel A of Table 4 and 5 cast doubt on weak-form market efficiency and support the notion that moving average trading rules contain substantial information to predict changes in the Italian stock index. Panel B of Table 5

reports the results of the AR moving average trading rules for three sub-periods. For the first two sub-periods, all three strategies significantly beat the 1-day unconditional mean. For the last sub-period, only strategy 3 significantly beats the buy-and-hold strategy. Again, it seems that the Italian equity market has become more efficient in recent years.

In conclusion, we have identified a strategy (strategy 3 based on AR moving average rules) that is based on technical trading and beats significantly the buy-and-hold strategy over the entire 30-year period and all three sub-periods. Although this discovery of profitable trading rules may be helpful in understanding market dynamics, traders may not be able to exploit these rules without considering transaction costs. In order to account for transaction costs of strategy 3 (leverage/money) based on AR trading rules, we report in Table 6, the "break-even" transaction costs, which are the percentage one-way costs that eliminate the additional return from technical trading. We also report annual transaction costs assuming that one-way transaction cost is .25 %.

In the first column of Table 6, we identify AR trading rules; in the second column we copy from Table 5 the average daily difference returns between the strategy 3 (leverage/money) and the buy-and-hold strategy. In column three, we obtain the annualized excess returns by multiplying average daily differences by 250 trading days (it is assumed that the market is closed 10 days per year). Technical trading rules imply that a trader be in the market when the rule emits a buy signal and be out of the market when the rule emits a sell signal; in column 4 we report total trades, numbers of in and out of the market signals, or total frequency of transactions, implied by a specific rule. We divide total trades by the number of years under consideration to get average trades per year. In order to estimate one-way "break-even" transaction cost, we divide annual excess return by average annual trade per year. One-way transaction cost is assumed to be the same for buying and selling the index. Finally, in the last column we estimate annual transaction costs assuming that one-way transaction cost is .25 %. Given that Knez and Ready (1996) have estimated one-way transaction cost for DJIA index to be .13%, and that Bessembinder and Chan estimate a one-way equity trading cost of .24% to .26% for institutional traders and that Wells Fargo Nikko Investment Advisors estimates its own one-way trading costs to be .25% (Kroner 1995), we think that our choice of .25% one-way transaction cost seems to be close to the real life situation.

Table 6. Strategy 3 based on AR moving average rules and break-even trading costs

Panel A: 1973-2002						
Rules	Strategy 3 $X(ddif)$	Excess return %	Total trades	Trades per year	Break-even costs %	Annual costs %, cost per trade =.25 %
$P > MA(5,50)$	0.00121	30.25	1123	37.43	.81	9.36
$P > MA(5,100)$	0.00111	27.75	1094	36.47	.76	9.12
$P > MA(5,150)$	0.00104	26.00	1069	35.63	.73	8.91
$P > MA(5,200)$	0.00102	25.50	1071	35.70	.71	8.93
$P > MA(10,50)$	0.00104	26.00	803	26.77	.97	6.69
$P > MA(10,100)$	0.00091	22.75	758	25.27	.90	6.32
$P > MA(10,150)$	0.00084	21.00	739	24.63	.85	6.16
$P > MA(10,200)$	0.00086	21.50	736	24.53	.88	6.13
Average	0.00100	25.09	924.13	30.80	.83	7.70
Panel B: sub-periods						
73-82: $P > MA(5,50)$	0.00152	38.00	381	38.10	1.00	9.53
83-92: $P > MA(5,50)$	0.00136	34.00	331	33.10	1.03	8.28
93-02: $P > MA(5,50)$	0.00077	19.25	441	44.10	.44	11.03

$X(ddif)$ is average daily difference between the return of strategy 3 (leverage/money) and the buy-and-hold strategy from Table 5. Excess return (annual) is $X(ddif)$ times 250. Total trades represent the number of total trades or total switch from in and out of the market. Trade per year is total trades divided by number of years. Break-even costs are estimated by the ratio of annual excess return over trades per year. Annual costs assuming one-way transaction cost to be .25 % is number of trades per year times .25%.

Table 6 shows that the break-even one-way trading costs, which would just eliminate excess returns from using the AR technical trading rules, are .83 % for the full sample and 1.00%, 1.03%, and .44% for 1973-1982, 1983-1992, and 1993-2002 sub-periods respectively. We believe that these break-even one-way trading costs are large compared to recent estimates of actual trading costs (.25%). If a trader's one-way transaction cost is below the break-even transaction costs of Table 6, the trader can use the above strategy and beat the buy-and-hold strategy even considering transaction costs. Table 6 also reports average annual costs of trading in and out of the market assuming a one-way trading cost of .25%; these annual trading costs are much lower than excess return from using trading rules, implying profitable technical trading for the Italian stock market even when considering transaction costs. In summary, our results provide support for technical trading rules that can be exploited.

5. *Conclusions*

Several papers have recently presented evidence that some simple trading rules are useful for predicting stock market returns. In this paper we investigate two moving average trading rules for the Italian stock index (Datastream index) over the period 1973-2002. Overall our results provide strong support for the technical trading rules that we explored. If technical analysis does not have any power to forecast price movements, then we should observe that the buy days returns do not differ appreciably from sell days returns; our results show that all the buy-sell differences are positive and the *t*-stats for these differences are highly significant, rejecting the null hypothesis of equality of buy days returns with sell days returns. The *t*-values for most buy and most sell are highly significant, rejecting the null hypothesis that the buy and sell returns equal the unconditional 1-day return (Tables 2 and 3). Our results indicate that moving average rules do indeed have predictive power and could discern recurring-price patterns for profitable trading. Given this predictive power of technical analysis, can we design various trading strategies to beat the buy-and-hold strategy? Our results on trading strategies support the hypothesis that technical trading rules can outperform the buy-and-hold strategy. Strategy 3 (the trader will borrow at the money market rate and double stock investment when trading rules emit buy signals and be in the money market when it emits sell signals) beats the-buy-and-hold strategy over the full sample and in each sub-period. Strategy 1 (the trader will be in the stock market when trading rules emit buy signals and be in the money market when it emits sell signals) beats the market over the full sample and two out of three sub-periods. Although the discovery of profitable trading rules may be helpful in understanding market dynamics, traders must also consider transaction costs. For strategy 3 that beats the buy-and-hold strategy over the entire 30-year period and all three sub-periods, we have estimated the break-even one-way trading costs that we believe are large compared to recent estimates of actual trading costs (.25 %). We conclude that technical trading rules have predictive power and that we could design a trading strategy that will beat the buy-and-hold strategy for the Italian equity index returns. The observation that technical trading rules have predictive power for changes in the Italian stock market index is consistent with weak-form market inefficiency.

REFERENCES

ALEXANDER S., "Price Movements in Speculative Markets: Trends or Random Walks", *Industrial Management Review*, 1964, 5(*2*), pp. 25-46.

ARNOLD C.M. and RAHFELDT, D., *Timing the Market: How to Profit in Bull and Bear Markets with Technical Analysis*, Chicago: Probus Publishing, 1986.

BESSEMBINDER H. and CHAN K., "The Profitability of Technical Trading Rules in the Asian Stock Markets", *Pacific-Basin Finance Journal*, 1995, *3*, pp. 257-84.

—— and ——, "Market Efficiency and Returns to Technical Analysis", *Financial Management*, 1998, *27*(2), pp. 5-17.

BROCK W., LAKONISHOK J., and LEBARON B., "Simple Technical Trading Rules and the Stochastic Properties of Stock Returns", *Journal of Finance*, 1992, *47*, pp. 1731-64.

ERGUL A., HOLMES P., and PRIESTLEY R., "Technical Analysis, Trading Volume and Market Efficiency: Evidence from an Emerging Market", *Applied Financial Economics*, 1997, *7*, pp. 361-65.

FAMA E., "The Behavior of Stock Market Prices", *Journal of Business*, 1965, *38*, pp. 34-105.

——, "Efficient Capital Markets: A Review of Theory and Empirical Work", *Journal of Finance*, 1970, *25*, pp. 383-417.

—— and BLUME M., "Filter Rules and Stock Market Trading Profits", *Journal of Business*, 1966, *39*, pp. 226-341.

GENCAY R. (1998a), "Optimization of Technical Trading Strategies and Profitability in Security Markets", *Economics Letters*, 1998, *59*(2), pp. 249-54.

—— (1998b), "The Predictability of Security Returns with Simple Technical Trading Rules", *Journal of Empirical Finance*, 1998, *5*(4), pp. 347-59.

GRANGER C. and MORGENSTERN O., "Spectral Analysis of New York Stock Market Prices", *Kyklos*, 1963, *16*, pp. 1-27.

JENSEN M. and BENINGTON G., "Random Walks and Technical Theories: Some Additional Evidence", *Journal of Finance*, 1970, *25*, pp. 469-82.

KNEZ P.J. and READY M.J., "Estimating the Profits from Trading Strategies", *Review of Financial Studies*, 1996, *9*, pp. 1121-64.

KRONER K., "Comments on: Do the Profits from Technical Trading Rules Reflect Inefficiencies?", mimeo, Wells Fargo Nikko Investment Advisors Advanced Strategy Group, San Francisco, 1995.

KWON K. and KISH R., "Technical Trading Strategies and Return Predictability: NYSE", *Applied Financial Economics*, 2002, *12*, pp. 639-53.

LARSON A., "Measurement of Random Process in Futures Prices", *Food Research Institute*, 1960, *1*, pp. 313-24.

MAILLET B. and MICHEL T., "Further Insights on the Puzzle of Technical Analysis Profitability", *European Journal of Finance*, 2000, *6*, pp. 196-224.

MANDELBROT B., "The Variation of Certain Speculative Prices", *Journal of Business*, 1963, *36*, pp. 394-419.

OSBORNE M., "Periodic Structure in the Brownian Motion of Stock Prices", *Operations Research*, 1962, *10*, pp. 345-79.

PRING M.J., *Technical Analysis: Explained*, New York: McGraw-Hill Co., 1991.

PRUITT S. and WHITE R., "The CRISMA Trading System: Who Says Technical Analysis Can't Beat the Market?", *Journal of Portfolio Management*, 1998, *14*(3), pp. 55-58.

RATNER M. and LEAL R.P.C., "Test of Technical Trading Strategies in the Emerging Equity Markets of Latin America and Asia", *Journal of Banking and Finance*, 1999, *23*(1), pp. 1887-905.

READY M., "Profits from Technical Trading Rules", *Financial Management*, Autumn 2002, pp. 43-61.

SHLEIFER A., *Inefficient Markets: An Introduction to Behavioral Finance*, Oxford: Oxford University Press, 2000.

TAYLOR M. and ALLEN H., "The Use of Technical Analysis in the Foreign Exchange Market", *Journal of International Money and Finance*, 1992, *11*, pp. 304-14.

VAN HORN J.C. and PARKER G.C., "The Random Walk Theory: An Empirical Test", *Financial Analyst Journal*, 1967, *23*, pp. 87-92.

ABSTRACT

This paper tests two moving average technical trading rules for the Italian stock market. Our results indicate that moving average rules do indeed have predictive power and could discern recurring-price patterns for profitable trading. Moreover, our results support the hypothesis that technical trading rules can outperform the buy-and-hold strategy. Break-even one-way trading costs are estimated to be .83% for the full sample and 1.00%, 1.03%, and .44% for 1973-1982, 1983-1992, and 1993-2002 sub-periods respectively. We believe these break-even costs are large compared to recent estimates of actual trading costs, implying profitable trading rules for the Italian stock market.

JEL classification: *C22, F31, G15*
Keywords: *technical trading rules, Italian stock market*

RISEC, Volume 50 (2003), No. 4, 451-474

THE DEMAND FOR MONEY BY PRIVATE FIRMS IN A REGULATED ECONOMY: THEORETICAL UNDERPINNINGS AND EMPIRICAL EVIDENCE FOR GERMANY, 1960-1998

by
MARTIN T. BOHL* and FRIEDRICH L. SELL**

1. *Introduction*

While empirical investigations into the properties of money demand have concentrated on the aggregate level, the famous 'finance motive' of money holdings put forward by Keynes (1936) and the early theoretical investigations by Baumol (1952) and Miller and Orr (1966) using an inventory-theoretic framework seem to point more at the business sector than at the aggregate level. More recently, there have been only a few remarkable studies on the money demand of firms (Fase and Winder, 1990; Barr and Cuthbertson, 1992; Mizen, 1996; Viren, 1996; Mulligan, 1997). As it appears to us, none of the authors mentioned, however, has attempted to link money demand of the business sector to the demand for labor on informal, likewise grey labor markets in the shadow economy.

* Department of Economics, European University Viadrina Frankfurt/Oder, Frankfurt/Oder (Germany). E-mail: bohl@euv-frankfurt-o.de

** Corresponding author. Department of Economics, Munich University of the Federal Armed Forces, Neubiberg (Germany). E-mail: friedrich.sell@unibw-muenchen.de

We thank Axel Jochem and Jochen Michaelis for their invaluable comments on Section 2 and participants of the session "Labor Market Economics" at the 47th conference of the International Atlantic Economic Society, March 16-23, 1999 in Vienna for their suggestions to an earlier version of this paper. An anonymous referee of this journal made significant contributions and helped to clear a number of crucial points in the paper. The usual disclaimer applies.

It is a matter of daily observation in highly industrialized economies that firms are inclined to hire work from informal markets whenever the burden of labor costs – in addition to the wage rate before taxes – makes this alternative profitable vis-à-vis the formal labor market. The sector of housing and construction is a typical and perhaps the most prominent example where companies draw heavily on labor input from informal sources and do hold cash exactly for this purpose. We are not so much interested – as some of the mentioned papers are – in the existence of economies of scale in the demand for cash by firms. Moreover, our main focus is on the determinants of money holdings by firms under the conditions of a highly regulated economy, especially in the labor market.

In glancing through the literature, it turned out that for the attempt to model firms' money demand in such an environment, much can be learned from a body of literature interested in – at first sight – quite a different issue, namely the demand for money in an economy with various constraints (Lane 1990a, 1990b, 1992). In these papers, Lane models households as choosing optimal money holdings subject to constraints in formal and informal goods as well as foreign exchange markets during the socialist experiment. Also, Lane incorporates the precautionary motive of money holdings into his analysis. In the firms' view of our own work, this can be translated into a precautionary finance motive on the background of informal labor supply. The cash-in-advance approach chosen by Lane will be used, here, as well. Recently, Bohl and Sell (1998) have done a similar methodological exercise for the demand for currency in an open economy.

The paper is organized as follows. In the following Section 2, our own model will be presented. In Section 3, the derived money demand function is subject to an empirical test with data from Germany applying modern econometric techniques. Thereafter, we will give a brief review on the empirical results achieved by other studies on the money demand of firms in Section 4. Finally, in Section 5, we give a summary of our results and some hints as to future research necessities.

2. *The Model*

The following analysis is not only a full application, but also an extension of Lane's (1992) cash-in-advance model for the household money demand in Poland to firms' money demand in a market economy with a significant shadow economy. A representative firm maximizes expected output, which is a function of labor input, over an infinite time horizon. The firm holds two

cash balances, both in domestic money. Purchases of labor input must be backed with money. Apparently, this assumption only applies to a minority of sectors in a modern economy and, in the past, it could be most likely found in the services and/or in the construction sector. For simplicity, it is assumed that labor is the only factor of production. The price for labor is lower in the shadow economy, and the quantity available there is unlimited very much in the vein of Lewis' (1954) seminal paper, whereas in the official market, the firm may not hire labor as much as desired, and, if so, only at a higher price. Therefore, domestic currency is held with the view to hire labor at a favorable price on the parallel labor market.[1]

The timing of transactions in the model is as follows. In the morning, the firm hires workers in the two markets, constrained by money held over from the previous day. In the afternoon, the firm's production is sold. There are no other financial assets. This assumption keeps the following algebra tractable and, at the same time, corresponds to the fact that 'money' is one of the few assets which can be observed both in the legal and in the shadow economy. The price and the available quantity of labor in the shadow economy, as well as the price of labor in the official market, are treated as random variables, whose realization is not known until the beginning of the day.

The firm maximizes expected output y, which is a function of factor input, i, over an infinite horizon:

$$EO = \bar{E}_0 \sum_{t=0}^{\infty} \beta^t y(i_t) \rightarrow \text{Max!} \tag{1}$$

where E_t denotes the expectations operator conditional on information available at time t. β is a discount factor (hence, an implicit interest rate is involved) and EO stands for expected output. For the sake of simplicity, factors of production only consist of labor, which is hired either on official or on parallel markets:

$$i_t = f_t^o + f_t^s \tag{2}$$

where f_t^o and f_t^s is the amount of labor hired in the official market and the shadow economy, respectively. Maximizing output, as stated in equation (1), is, according to duality theory (Nicholson, 1992), equivalent to the minimization of costs for a given (but not necessarily constant over time)

[1] See Sell (1997) for an analysis of the income velocity of money in the informal sector.

total cost of inputs ($TC_t = f_t^o l_t^o + f_t^s l_t^s$). Insofar, our approach is capable to explore the optimal production plan of the firm.[2] In addition, we have a view on the optimal financial plan of the firm: The firm's financial budget constraint for each period requires that expenses for labor during the period plus end-of-period money holdings are financed by money carried over from the previous period and by sold output[3]:

$$f_t^o l_t^o + f_t^s l_t^s + m_t^{d_o} + m_t^{d_s} = m_{t-1}^{d_o} + m_{t-1}^{d_s} + y_t \tag{3}$$

where l_t^o denotes the price of labor on the official market, l_t^s the price of labor in the shadow economy, $m_t^{d_o}$ the money holdings for the official market and $m_t^{d_s}$ the money holdings for the shadow market. In an intertemporal equilibrium context, equation (3) should incorporate the 'zero net profits' condition for the competitive firm:

$$\sum_{t=0}^{\infty} \beta^t \left[y_t - (f_t^o l_t^o - f_t^s l_t^s) \right] = 0 \Leftrightarrow \sum_{t=0}^{\infty} \beta^t \left[(m_t^{d_o} - m_{t-1}^{d_o}) + (m_t^{d_s} - m_{t-1}^{d_s}) \right] = 0 \tag{3a}$$

At first glance, the distinction between two different cash holdings – one legal and one illicit – of the same currency by a representative firm may seem to be artificial. But one should be aware of the fact that legal cash belongs to the official assets of the firm and, hence, enters the firm's books while illicit cash never does. Moreover, in making the distinction between these two components of currency demand, we follow the procedure of Bhattacharyya (1990). As opposed to Bhattacharyya, however, here both components and their economic determinants will be derived from optimality conditions in a constrained maximisation problem. Expenses for hiring labor in the shadow economy are limited by holdings of illicit cash balances carried over from the previous period:

$$f_t^s l_t^s \leq m_{t-1}^{d_s} \tag{4}$$

Previous money holdings for the shadow labor market is a predetermined

[2] As money holdings help the firm to hire labor both from official and from parallel markets, the budget constraint for each period relates actual (end-of-period) money holdings and cash carried over from the previous period (both stock variables) to flow variables such as sold output or expenses for labor.

[3] For a constant price level of 1, production equals sold output: $p_t y_t = Y_t$ for $p_t = 1$.

variable in the model.[4] Total expenses for hiring labor on the two markets are constrained by the two types of cash balances:

$$f_t^o l_t^o + f_t^s l_t^s \leq m_{t-1}^{d_o} + m_{t-1}^{d_s} \tag{5}$$

The amount of labor that the firm can hire in the shadow economy is not limited by the availability of supply and hence it is always lower than the available amount. Therefore, we assume that the representative firm faces infinitely elastic input supplies in the shadow economy. In the official labor market, however, labor supply is limited to the amount $\bar{f}_t^A$. Hence:

$$f_t^o \leq \bar{f}_t^A \tag{6}$$

where the quantity of labor available is $\bar{f}_t^A$. There are also nonnegativity conditions on the holdings of the two cash balances and on the hiring of labor in the two markets[5]:

$$m_t^{d_o}, m_t^{d_s} \geq 0 \tag{7}$$

and:

$$f_t^o, f_t^s \geq 0 \tag{8}$$

The firm, therefore, maximizes the expected discounted value of production (equation (1)) with respect to the quantities of hired labor (in each market) and to the holdings of legal and illicit cash in each period, subject to the constraints (3) through (8). Defining a value function, $J(m_t^{d_o}, m_t^{d_s})$, as the expected maximized present value of production from period $(t + 1)$ onward as a function of money carried over from period t, the Lagrangean for each period t can be written as follows:[6]

---

[4] There are a number of possibilities to endogenize the process of replenishing illicit cash balances. For instance, we could argue that output (see equation (3) above) can be sold on official as well as on parallel markets, the latter implying a good's price free of taxes. The revenue on the parallel goods market, then, would serve to replenish illicit cash in order to hire labor on the shadow labor market in the next period.

[5] All in all, there are three sources for domestic demand for cash balances: legal and illicit cash of firms, cash balances of private households, and, finally, demand for domestic cash from abroad $m_t^d = m_t^{d_o} + m_t^{d_s} + m_t^{d_{hh}} + m_t^{d_f}$.

[6] Notice: this formulation of the constrained maximisation implies that $\lambda_{i,t}$ $(i = 1, 2, 3, 4)$

$$
\begin{aligned}
(9) \qquad L = {} & y\left(f_t^o + f_t^s\right) + \beta J\left(m_t^{d_o}, m_{t_1}^{d_s}\right) + \lambda_{0,t}\left(TC - f_t^0 l_t^0 - f_t^s l_t^s\right) \\
& - \lambda_{1,t}\left(f_t^o l_t^o + f_t^s l_t^s + m_t^{d_o} + m_t^{d_s} - m_{t-1}^{d_o} - m_{t-1}^{d_s} - y_t\right) \\
& - \lambda_{2,t}\left(f_t^s l_t^s - m_{t-1}^{d_s}\right) - \lambda_{3,t}\left(f_t^o l_t^o + f_t^s l_t^s - m_{t-1}^{d_o} - m_{t-1}^{d_s}\right) \\
& - \lambda_{4,t}\left(f_t^o - \bar{f}_t^A\right) - \lambda_{5,t} m_t^{d_o} - \lambda_{6,t} m_t^{d_s} - \lambda_{7,t} f_t^o - \lambda_{8,t} f_t^s
\end{aligned}
$$

Now, we define:

$$
(10) \qquad \theta_t = \frac{l_t^o - l_t^s}{l_t^o} > 0
$$

where θ_t is the excess price for labor to be payed by entrepreneurs in the official market relative to the market in the shadow economy. We shall call it the 'wage wedge' hereafter. By this we mean (as a percentage rate) the wedge between labor costs and net income of employees (take home pay). The price which accrues from hiring labor in the official market can be calculated according to the compensation of employees concept (Hinze, 2000). In this concept it is assumed that firms have to afford – in excess of gross wages and salaries – the respective social security payroll tax. Opposed to this, the employees' utility reservation wage level corresponds to the net income of employees (wages/salaries) after deduction for personal income tax and social security contributions. The price for hiring labor in the shadow economy equals in equilibrium by and large the utility reservation wage level from the official economy.

Using this device, one obtains the following *first order conditions for hiring labor* in the two markets:

$$
(11a) \qquad \frac{\partial y}{\partial f_t^o} = l_t^o\left(\lambda_{0,t} + \lambda_{1,t} + \lambda_{3,t}\right) + \lambda_{4,t} + \lambda_{7,t} = \frac{\partial y}{\partial i_t} = y'_t
$$

and:

$$
(11b) \qquad \frac{\partial y}{\partial f_t^s} = l_t^s\left(\lambda_{0,t} + \lambda_{1,t} + \lambda_{2,t} + \lambda_{3,t}\right) + \lambda_{8,t} = \frac{\partial y}{\partial i_t} = y'_t
$$

where the $\lambda_{i,t}$ (i = 1,...,8) are static Lagrangean multipliers (see also equation (2)). Equations (11a) and (11b) are the first order conditions for hiring labor

---

are positive and $\lambda_{i,t}$ (i = 5, 6, 7, 8) are negative. See Kreps (1994, p. 704).

in the two markets. As in Lane (1992) we assume, for simplicity, that some labor is hired in each of the two labor markets in each period so that $\lambda_{7,t} = \lambda_{8,t} = 0$. Otherwise, the tax authorities could anticipate a production in the shadow economy. Hence, conditions (11a) and (11b) can be shortened to become:

(11a') $$\frac{\partial y}{\partial i_t} = y'_t = l_t^o(\lambda_{0,t} + \lambda_{1,t} + \lambda_{3,t}) + \lambda_{4,t}$$

and:

(11b') $$\frac{\partial y}{\partial i_t} = y'_t = l_t^s(\lambda_{0,t} + \lambda_{1,t} + \lambda_{2,t} + \lambda_{3,t})$$

Conditions (11a') und (11b') can be combined to yield:

(12) $$y'_t = l_t^s(\lambda_{0,t} + \lambda_{1,t} + \lambda_{2,t} + \lambda_{3,t}) = l_t^o(\lambda_{0,t} + \lambda_{1,t} + \lambda_{3,t}) + \lambda_{4,t}$$

which, solved for $\lambda_{2,t}$, gives:

(12') $$\lambda_{2,t} = \frac{l_t^o - l_t^s}{l_t^s}(\lambda_{0,t} + \lambda_{1,t} + \lambda_{3,t}) + \frac{\lambda_{4,t}}{l_t^s}$$

Introducing (12') into (11b'), we get:

(13) $$y'_t = l_t^s\left(\lambda_{0,t} + \lambda_{1,t} + \frac{l_t^o - l_t^s}{l_t^s}(\lambda_{0,t} + \lambda_{1,t} + \lambda_{3,t}) - \frac{\lambda_{4,t}}{l_t^s} + \lambda_{3,t}\right)$$

Rearranging (12') gives (12a):

(12a) $$\left(\lambda_{2,t} - \frac{\lambda_{4,t}}{l_t^s}\right)\left(\frac{l_t^s}{l_t^o - l_t^s}\right) = (\lambda_{0,t} + \lambda_{1,t} + \lambda_{3,t})$$

Then, (12a) introduced into (11a') gives:

(14) $$y'_t = \left(\lambda_{2,t} - \frac{\lambda_{4,t}}{l_t^s}\right)\left(\frac{l_t^s}{l_t^o - l_t^s}\right) l_t^o$$

or:

$$y'_t = \left(\lambda_{2,t} l_t^o - \lambda_{4,t} \frac{l_t^o}{l_t^s}\right)\left(\frac{l_t^s}{l_t^o - l_t^s}\right) \tag{14'}$$

or, finally:

$$y'_t = \left[\lambda_{2,t} l_t^o - \lambda_{4,t}\left(1 + \frac{l_t^o - l_t^s}{l_t^s}\right)\right]\left(\frac{l_t^s}{l_t^o - l_t^s}\right) \tag{14''}$$

Because of:

$$\frac{\theta_t}{(1-\theta_t)} = \frac{l_t^o - l_t^s}{l_t^s} \tag{1θa}$$

we can rewrite (14'') to become:

$$y'_t = \left[\lambda_{2,t} l_t^o - \lambda_{4,t}\left(1 + \frac{\theta_t}{1-\theta_t}\right)\right]\left(\frac{1-\theta_t}{\theta_t}\right) \tag{14a}$$

Rearranging (14a) leads to (14b):

$$\frac{\theta_t}{1-\theta_t} = \frac{\lambda_{2,t} l_t^o - \lambda_{4,t}}{y'_t + \lambda_{4,t}} \tag{14b}$$

The wedge between the prices of labor in the official and in the parallel market is either associated with shortages in the official labor market ($\lambda_{4,t}$) or with a lack of money to pay labor on the parallel labor market ($\lambda_{2,t}$). The rationale of equation (14b) is straightforward: Whenever the constraint of supply on the official labor market becomes more binding, additional demand for labor spills over to the informal labor market and narrows the wage wedge between the two markets. In other words, given the low prices and unlimited quantities in the parallel labor market, firms either hire everyone who is available at the low wage costs (still taking into account the non-negativity condition (8)), or run out of illicit cash in the attempt.

Next[7], we consider the first order conditions for optimal holdings of legal and illicit cash:

$$\frac{\partial L}{\partial m_t^{d_0}} = \beta \frac{\partial J}{\partial m_t^{d_0}} - \lambda_{1,t} - \lambda_{5,t} = 0 \tag{15}$$

$$\beta J_{d_0,t} = \lambda_{1,t} + \lambda_{5,t} \tag{15a}$$

$$\frac{\partial L}{\partial m_t^{d_s}} = \beta \frac{\partial J}{\partial m_t^{d_s}} - \lambda_{1,t} - \lambda_{6,t} = 0 \tag{16}$$

$$\beta J_{d_s,t} = \lambda_{1,t} + \lambda_{6,t} \tag{16a}$$

where:

$$J_{i,t} = \frac{\partial J(m_t^{d_0}, m_t^{d_s})}{\partial m^{i,t}} \qquad i = d_0, d_s$$

Using the following definition of the value function, we can assess the contributions of money holdings to future production (see also the Lagrangean above):

$$J_{d_0,t} = \frac{\partial J}{\partial m_t^{d_0}} = E_t\left(\lambda_{1,t+1} + \lambda_{3,t+1}\right) \tag{17}$$

$$J_{d_s,t} = \frac{\partial J}{\partial m_t^{d_s}} = E_t\left(\lambda_{1,t+1} + \lambda_{2,t+1} + \lambda_{3,t+1}\right) \tag{18}$$

That is, legal cash is expected to contribute to future production by easing the budget constraint (reflected in $\lambda_{1,t}$) as well as to the extent that it may ease a cash-in-advance constraint ($\lambda_{3,t}$). Holding illicit cash is expected to ease the budget constraint ($\lambda_{1,t}$) and two cash-in-advance constraints ($\lambda_{2,t}$, $\lambda_{3,t}$).[8] In the mathematical annex, we show how the four mentioned first order conditions

[7] Hence, the solution of the optimization problem is equal to the determination of the optimal inputs of labor and the optimal provision of (legal and illicit) cash as *instruments* to hire labor and to secure (future) production.

[8] Both types of money holdings are, hence, considered as *instrumental variables* (in the economic and not in the econometric sense) and not at the same time as parametric variables in the model.

can be combined to yield a proxy for the determinants of illegal and legal cash. Namely, in equation (28) – see Annex – we find a formulation for the relative attractiveness to hold illicit rather than legal cash.

Equation (29)[9], which is a second-order Taylor approximation of equation (28), can then be used to provide a portfolio-theoretic rationale for the (il)legal demand for domestic currency. First, the demand for (il)legal cash depends (positively) negatively on the expected excess price of labor in the official market over the price of labor in the parallel market. This labor price premium is divided by the expected parallel market wage increase factor, which deflates the expected labor price premium. Second, the higher the parallel market wage increase factor, the (lower) higher is the demand for (il)legal cash, ceteris paribus. Demand for (il)legal cash depends on the variances and covariances of labor input and wage inflation on the parallel market. The implication of the variance of labor input is ambiguous and depends on the third derivative of the production function y'''. Third, a higher variance of wage inflation on the parallel market is associated with a (higher) lower demand for (il)legal cash. Also, the demand for (il)legal cash is (negatively) positively related to the covariance between expected labor input and the expected wage premium of the official over the parallel market. In the same vein, the demand for (il)legal cash is (positively) negatively associated with the covariance between expected labor input and wage inflation on the parallel market. Fourth, the demand for (il)legal cash is (negatively) positively related to the covariance between wage inflation on the parallel market and the wage premium of the official over the parallel market. Taking into consideration the role of output in equation (3) and the fact that labor input is not an observable variable, the demand for (il)legal cash can be expressed as a function of the following five determinants:

$$(3\theta) \qquad m_t^{d_a} = m\left(\overset{+}{y_t}, \overset{-}{E_t\theta_{t+1}}, \overset{+}{E_t\pi_t^s}, \overset{-}{\operatorname{var}\pi_t^s}, \overset{+}{\operatorname{cov}(\theta_{t+1}, \pi_t^s)} \right)$$

Hence, mutatis mutandis demand for illicit cash reads:

$$(31) \qquad m_t^{d_x} = m\left(\overset{+}{y_t}, \overset{+}{E_t\theta_{t+1}}, \overset{-}{E_t\pi_t^s}, \overset{+}{\operatorname{var}\pi_t^s}, \overset{-}{\operatorname{cov}(\theta_{t+1}, \pi_t^s)} \right)$$

It comes as no surprise that neither legal nor illicit cash depend on the interest rate or on any other opportunity cost variable – given that there are

[9] See the mathematical Annex for a full and step by step derivation of equation (29).

no other financial assets than money (see above) in the model.[10] As both components of currency demand are not observable individually, there seems to be a problem to estimate a currency demand of firms equation. As only production enters into both components/demand functions with the same sign and both demand functions are symmetric in all the other variables with regard to the sign, one way out would be to estimate a cash demand by firms function, where only production enters as a determinant. The disadvantage of doing so, however, would consist in losing all the additional information embedded in the components (30), (31). Hence, a second way out consists in building the aggregate as in Bhattacharyya (1990):

$$(32) \quad m_t = m_t^{d_1} + m_t^{d_2} = (\alpha_{d_1} + \alpha_{d_2}) y_t + (\beta_{d_1} - \beta_{d_2}) E_t \theta_{t+1} + (\gamma_{d_2} - \gamma_{d_1}) E_t \pi_t^s + (\delta_{d_1} - \delta_{d_2}) \operatorname{var} \pi_t^s + (\varepsilon_{d_2} - \varepsilon_{d_1}) \operatorname{cov}(\theta_{t+1}, \pi_t^s)$$

If the illicit (official) component dominates, then the estimated five compound coefficients in (32) should have the following signs: α, β, $\delta > 0$ (< 0), but γ, $\varepsilon < 0$ (> 0).

3. *Data and Empirical Results*

The time series used in our study are annual observations covering the period from 1960 to 1998. The monetary aggregate for the business sector consists of cash balances and sight deposits and is from the financial accounts for Germany, published by the Deutsche Bundesbank (Deutsche Bundesbank, 1994, 1999). The statistics of the Deutsche Bundesbank (Finanzierungsrechnung) do not allow for a breakdown of money holdings into cash and sight deposits (Brümmerhoff, 1985). Therefore, when estimating money demand functions of firms instead of demand for cash balances functions (Bohl and Sell, 1998) of these firms, additional motives of holding sight deposits, which go beyond the scope of this paper, come into play. The other problem – that a breakdown of the sectors', and hence also firms' cash holdings into their (its) different components is usually unavailable – is not a hindering factor in principle for our empirical analysis, as we have shown in the last section. Financial assets and liabilities of firms are derived from the statistics of financial institutions and not from the balance sheets of the firms (Bundesbank, 1999). This makes sure that we deal with net financial assets

[10] See for an analogous result Seitz (1995).

which are not those observed by the fiscal authorities and can be subject of payments to the shadow economy. The information on cash holdings of firms results from consistency rules applied to all gross and net balances for the non-financial sectors and should, hence, cover all (legal and illicit) cash transactions of this sector. We deflate this time series by the GDP deflator to construct real money balances m_t. The GDP deflator is taken from the annual report of the German Council of Economic Experts (Sachverständigenrat, 2001). The gross value added of the business sector in constant prices serves as the scale variable y_t.

What about the size of the shadow economy in Germany? Recent calculations by Schneider (1999) say that the income generated in the parallel economy at the beginning of the new century corresponds to approximately 15% of GDP. Empirical estimates with econometric techniques (Karmann, 1986, 1990) show that the share of the shadow economy was almost zero in Germany in 1970. Since then, it has been rising and reached a level of 10% of GNP already in 1990. If Schneider's recent calculations are correct, the speed of growth of the hidden economy has accelerated during the 1990s.

To construct the wage wedge w_t we divide the total cost per employee (according to the compensation of employee concept) by the net wage/salary per employee. This ratio was 1.376 in 1960 and has increased meanwhile (1999) to a level of approximately 1.934. Between 1960 and 1970 the ratio remained rather stable. Since 1970, however, the ratio has been ascending considerably. This finding matches nicely the behavior of the shadow economy (see above). Only since the end of the 1990s we do experience a relaxation in the speed of increase. The time series on which we build our empirical analysis are from the Hamburg Institute of International Economics (Hinze, 2000).

The variables m_t, y_t and w_t are in logarithms and refer from 1990 onwards to West and East Germany.[11] To use actual rather than expected values of the wage wedge, we rely on the argument put forward by Taylor (1991), according to which the application of cointegration techniques allows to test our money demand function subject only to the very weak assumption that forecasting errors are stationary.

[11] To take into account the German unification we have included in the regression equations a dummy variable.

Table 1. KPPS Tests

Panel A: KPSS tests for the null of level stationarity				
Variable	$l = 0$	$l = 1$	$l = 2$	$l = 3$
m_t	3.43*	1.84*	1.29*	1.02*
y_t	3.73*	1.95*	1.34*	1.04*
w_t	3.70*	1.93*	1.33*	1.04*
Panel B: KPSS tests for the null of trend stationarity				
Variable	$l = 0$	$l = 1$	$l = 2$	$l = 3$
m_t	0.53*	0.31*	0.23*	0.19*
y_t	0.56*	0.30*	0.22*	0.18*
w_t	0.49*	0.26*	0.19*	0.15*

Note: m_t is real money balances of firms, y_t real gross value added of the business sector and w_t the wage wedge. l are the truncation lags. * denotes significant statistics at the five per cent level where the critical values are from Kwiatkowski et al. (1992).

To investigate the degree of integration of the times series we employ the test suggested by Kwiatkowski et al. (1992), hereafter KPSS test. In the KPSS test the null hypothesis of the stationarity of a variable is tested against the alternative of a unit root. The results of these tests for the null hypothesis of level stationarity and trend stationarity employing the truncation lags from 0 to 3 are reported in Table 1 where the maximum lag length is chosen due to the suggestion in Schwert (1987) $l = \text{int}\{4(T/100)^{1/4}\}$. The KPSS tests are performed for the variables m_t, y_t and w_t assuming that the other variables in our money demand equation are stationary time series. As can be seen in Table 1, the null hypotheses of level and trend stationarity are rejected for all three time series because the test statistics are significant at the five per cent level. Hence, we assume that m_t, y_t and w_t are integrated of order one.

Next, we estimate and test for cointegration relying on the Johansen procedure (Johansen, 1988, 1991) between the non-stationary time series. Table 2 presents the tests of the null hypothesis of r cointegrating vectors using the trace test. The test statistics show the existence of a single cointegrating relationship because the null hypothesis of zero cointegrating vectors is rejected at the five per cent level while the hypothesis that the number of cointegrating vectors is less than or equal to one cannot be rejected. The underlying VAR model has the lag length one and the LM-type tests (LM1, LM4) show the absence of first and fourth order autocorrelation. When looking at the estimated cointegrating vector (after normalizing on m_t) the long-run income elasticity is close to one and the coefficient of the wage wedge has a positive sign.

Table 2. Johansen cointegration tests

Null	Trace	LM1	LM4	Estimated cointegrating vector
$r=0$	36.62*	10.45	4.82	$m_t = 0.98y_t + 2.52w_t$
$r\leq 1$	14.78			
$r\leq 2$	1.19			

Note: m_t is real money balances of firms, y_t real gross value added of the business sector and w_t the wage wedge. The VAR has the lag length one and includes a constant term. * denotes significant statistics at the five per cent level where critical values can be found in Osterwald-Lenum (1992).

Having analysed the stochastic properties of the individual time series and the long-run relationship, we have estimated an error correction model which inculdes all explanatory variables of our money demand function. Starting with this general specification we have excluded step by step the variables with statistically insignificant coefficients at the five per cent level. The result of this general-to-specific testing strategy is the following error correction model:

$$\Delta m_t = \underset{(2.56)^*}{0.06} + \underset{(2.66)^*}{1.47}\,\Delta w_{t-1} - \underset{(2.09)^*}{0.31}\,\hat{u}_{t-1} + \hat{\varepsilon}_t$$

$DW = 2.03 \quad Q_{(7)} = 4.99 \quad Q_{(3)} = 1.85 \quad \bar{R}^2 = 0.45 \quad RESET = 0.77$

Period: 1962 – 1998 $\quad u_t = m_t + 3.10 - 0.98y_t - 2.52w_t$

The model contains as explanatory arguments the stationary long-run relationship between real money, the scale variable and the wage wedge and in addition only the lagged wage wedge as an explanatory variable of the short-run dynamics. The parameter on the wage wedge has a positive sign. All other arguments of our short-run money demand model were statistically insignificant. Surprisingly, neither the variance and covariance terms of equation (30) and (31) nor the scale variable and a nominal interest rate measuring the opportunity cost of cash holdings were statistically significant.

This result is not contradictory to the implications of the model (see equations (30) and (31)). As the statistics of business cash balances and sight deposits are not calculated from the firms' books, but from sources of the consolidated monetary sector, they represent a real 'mix' of legal and illicit cash motives. A positive sign of the wage wedge in the above estimation equation, hence, is a signal for a significant (and dominant) influence of the parallel labor market on the money holdings of the firms. The coefficient of the error correction term is statistically significantly different from zero

which confirms the findings of the cointegration analyses. Furthermore, the residuals of the error correction model exhibit no autocorrelation as can be seen by the Durbin-Watson (DW) and Ljung-Box ($Q_{(df)}$) tests. In addition, the RESET test shows that there is no functional misspecification. Measured by the adjusted coefficient of determination ($\bar{R}^2$) the explanatory power of the model is acceptable. In summary, the error correction model has statistically and economically sensible characteristics and describes the behavior of the demand for money quite well.

4. *Findings of Related Studies and Comparison with Our Results*

As mentioned in the introduction, our study is part of the small literature on the money demand of firms so that it is sensible to review this branch of studies and compare their findings with our own results. Meltzer (1963), Whalen (1965) and Vogel and Maddala (1967) wrote the first articles investigating the behaviour of demand for money of companies and stimulated subsequent research activities. More recently, Fase and Winder (1990) specify error correction models to study aggregate and sectorally disaggregated money demand functions for M1 and M2 in the Netherlands over the period from 1970 to 1988. An income elasticity of unity was found for the business sector, the household and the aggregate money demand functions irrespective of the money definition used. The interest and inflation elasticities for the business sector are systematically higher in absolute terms than the values of the household sector money demand functions and the aggregate demand for money functions. Again, these findings are insensitive to the usage of different money definitions and could be a result of the greater scope which the business sector has for conducting active and systematic cash management compared with the household sector. In addition, Fase and Winder (1990) experiment with a cyclical indicator as part of the short-term dynamics and as a proxy for the costs of real assets for which the liquid financial assets are a substitute. The coefficients of this variable in all money demand functions for the business sector are statistically not significantly different from zero.

Barr and Cuthbertson (1992) analyse the company sector liquid asset holdings in the United Kingdom within a system framework for the period from 1976 to 1986. They implement cointegration techniques and obtain asset demand functions which satisfy the theoretical restrictions. The estimated demand functions are intuitively plausible and exhibit parameter stability. The demand of the company sector for liquid assets depends on

various rates of return and on wealth. In Mizen (1996), a forward-looking buffer stock model for the company sector in the United Kingdom over the period from 1970 to 1988 is specified. The results suggest that the buffer stock models for the money demand behaviour of the company sector are validated and that businesses hold money for the insulation against unanticipated shocks. Compared with the findings from studies on aggregate money demand functions the long-run cointegrating relationships reflect a different behaviour of the company sector, their money holdings are far more volatile and may well be interrelated with bank lending and net trade credit arrangements. These results are not surprising because the company sector actively monitors and adjusts its balances far more often than the household sector.

Among the recent investigations on the demand for money of companies there are two cross-sectional studies which should be mentioned. Mulligan (1997) provides a cross-sectional study of the demand for money relying on 12,000 firms in the United States for the period from 1961 to 1992. The money demand function considered relates money holdings to the volume of sales, the opportunity cost of holding money and the value of the cash manager's time. His findings indicate that there are economies of scale and that companies headquartered in countries with high wages hold more money for a given level of sales. The estimated scale as well as wage elasticities are by and large of equal amount and lie around 0.8 and the interest elasticity is statistically significantly negative. Viren (1997) analyses the relationship between the demand for cash and the transactions volume by 2,700 Finnish firms. Relying on this cross-section of business firms the findings show economies of scale and that the relationship of demand for cash and the transaction volume differs across the branches of the economy. Cash is not predominantly used in the shadow economy and is still a competitive means of payment in the service sector of the Finnish economy.

The review of the existing literature on the money demand of firms demonstrates that none of the studies has investigated theoretically nor empirically the importance of the demand for labor in the shadow market for the firms' money demand. The investigations rely mostly on the inventory theoretical model put forward by Baumol (1952) and Tobin (1956), while our own estimated money demand function is based on Lane's (1992) cash-in-advance model. The findings of the studies mentioned above can be reconciled with our own results concerning long-run elasticity of the scale variable because the estimated values are comparable. To our knowledge there is no investigation on the demand for money of firms for Germany. Virtually all studies analyse the aggregate money demand for M1 and M3

(see, for example, Falk and Funke, 1995; Hansen and Kim, 1995; Wolters et al., 1998). Comparing their findings with our results, it is interesting to note that the long-run elasticity of the scale variable is generally well above one, while the estimated long-run elasticity of the scale variable in the firms' money demand is near one. When looking at the error correction model the estimated parameter of the error correction term in our model is higher in absolute terms. This indicates faster adjustment towards equilibrium on the firms' level than on the aggregate level.

5. *Summary*

Based on Lane's (1992) cash-in-advance approach, the paper investigates theoretically the determinants of money holdings of firms under the conditions of a highly regulated labor market and analyses empirically the demand for money of German businesses during the period from 1960 to 1998. The paper differs from previous studies because none of these has attempted to link money demand of firms to the demand for labor in the shadow economy. Moreover, there are no empirical investigations concerning the demand for money of German firms.

As a result of our theoretical analysis the demand for cash balances by firms for shadow market activities depends among other things positively on the expected wage wedge, which is the excess price for labor to be paid by firms in the official market relative to the price to be paid in the shadow economy. Relying on this theoretical finding, we concentrate our interest on the importance of the wage wedge as an explanatory argument of a money demand function for German firms. The empirical results support the theoretical findings in the sense that the coefficient of the wage wedge has a positive sign in the long-run cointegrating relationship. Furthermore, the wage wedge appears with a statistically significant positive parameter in the short-run relationship of the error correction model, which captures the money demand behavior of German firms quite well. Hence, the results point at the relevance of illicit cash as a part of firms' money demand during the period under analysis.

MATHEMATICAL ANNEX

Equation (17) inserted into (15a) and (18) inserted into (16a) gives:

$$\beta E_t(\lambda_{1,t+1} + \lambda_{3,t+1}) = \lambda_{1,t} + \lambda_{5,t} \tag{19}$$

(20) $$\beta E_t(\lambda_{1,t+1} + \lambda_{2,t+1} + \lambda_{3,t+1}) = \lambda_{1,t} + \lambda_{6,t}$$

Equations (19) and (20) can be simplified by assuming that $\lambda_{3,t} = \lambda_{6,t} = 0$. In this case, illicit cash is also used as a long term store of value and the firm thus does not exhaust its illicit cash this period (later, we will proceed in the same way with regard to legal cash). Hence, neither the overall cash-in-advance constraint (equation (5)) nor the nonnegativity constraint on illicit cash balances (equation (7)) is binding in period t. Solving (11a') and (11b') for $\lambda_{1,t}$ (and for $\lambda_{1,t} + \lambda_{2,t}$, respectively) leads to (11c') and (11d') respectively:

(11c) $$y'_t = l_t^o(\lambda_{0,t} + \lambda_{1,t}) + \lambda_{4,t}$$

(11c') $$\frac{y'_t}{l_t^o} = \lambda_{0,t} + \lambda_{1,t} + \frac{\lambda_{4,t}}{l_t^o} \qquad \lambda_{1,t} = \frac{y'_t}{l_t^0} - \lambda_{0,t} - \frac{\lambda_{4,t}}{l_t^o}$$

(11d) $$y'_t = l_t^s(\lambda_{0,t} + \lambda_{1,t} + \lambda_{2,t})$$

(11d') $$\lambda_{0,t} + \lambda_{1,t} + \lambda_{2,t} = \frac{y'_t}{l_t^s} \qquad \lambda_{1,t} + \lambda_{2,t} = \frac{y'_t}{l_t^s} - \lambda_{0,t}$$

Applying (11d') to the next period and inserting into (20) gives:

(11d'') $$E_t(\lambda_{1,t+1} + \lambda_{2,t+1}) = E_t\left(\frac{y'_{t+1}}{l_{t+1}^s}\right) - E_t(\lambda_{0,t+1})$$

and, extending to:

(11d''') $$\beta E_t(\lambda_{1,t+1} + \lambda_{2,t+1}) = \beta E_t\left(\frac{y'_{t+1}}{l_{t+1}^s}\right) - \beta E_t(\lambda_{0,t+1})$$

Hence,

(21) $$\beta E_t\left(\frac{y'_{t+1}}{l_{t+1}^s}\right) - \beta E_t(\lambda_{0,t+1}) = \lambda_{1,t}$$

Introducing (11c') into (21) yields:

(22) $$\beta E_t\left(\frac{y'_{t+1}}{l_{t+1}^s}\right) - \beta E_t(\lambda_{0,t+1}) = \frac{y'_t}{l_t^o} - \frac{l_t^o\lambda_{0,t}}{l_t^o} - \frac{l_t^o\lambda_{4,t}}{l_t^o}$$

or:

(22') $$\beta E_t\left(\frac{y'_{t+1}}{l_{t+1}^s / l_t^o}\right) - \frac{\beta E_t(\lambda_{0,t+1})}{1/l_t^o} = y'_t - l_t^o\lambda_{0,t} - l_t^o\lambda_{4,t}$$

From (10), it follows:

(10') $$l_t^o = \frac{l_t^s}{1-\theta_t}$$

Introducing (10') into (22') gives:

(23) $$\beta E_t\left(\frac{y'_{t+1}}{(l^s_{t+1}/l^s_t)(1-\theta_t)}\right) - \frac{\beta E_t(\lambda_{0,t+1})}{(1/l^s_t)(1-\theta_t)} + \frac{l^s_t}{(1-\theta_t)}\lambda_{0,t} + \frac{l^s_t}{(1-\theta_t)}\lambda_{4,t} = y'_t$$

Define:

(24) $$\frac{l^s_{t+1} - l^s_t}{l^s_t} = \pi^s_t$$

and introduce into (23):

(25)

$$y'_t = \beta E_t\left(\frac{y'_{t+1}}{(1+\pi^s_t)(1-\theta_t)}\right) - \frac{\beta E_t(\lambda_{0,t+1})}{(\pi^s_t+1)/l^s_{t+1}\cdot(1-\theta_t)} + \frac{l^s_{t+1}/(\pi^s_t+1)}{(1-\theta_t)}\lambda_{0,t} + \frac{l^s_{t+1}/(\pi^s_t+1)}{(1-\theta_t)}\lambda_{4,t}$$

Alternatively, define:

(10b) $$\vartheta_t = \frac{l^o_t - l^s_t}{l^s_t}$$

(10b') $$\theta_t = \frac{\vartheta_t}{1+\vartheta_t}$$

Insert (10b') into (25) to get:

(25a) $$y'_t = \beta E_t\left(\frac{(y'_{t+1})(1+\vartheta_t)}{(1+\pi^s_t)}\right) - \frac{\beta E_t(\lambda_{0,t+1})(1+\vartheta_t)}{(\pi^s_t+1)/l^s_{t+1}} + \frac{\lambda_{0,t}l^s_{t+1}(1+\vartheta_t)}{\pi^s_{t+1}} + \frac{\lambda_{4,t}l^s_{t+1}(1+\vartheta_t)}{\pi^s_{t+1}}$$

Next, the condition for optimal holdings of legal cash can be written as:

(19a) $$\beta E_t(\lambda_{1,t+1}) = \lambda_{1,t} + \lambda_{5,t}$$

Applying (11c') to the next period and inserting (21) yields:

(26) $$\beta E_t\left(\frac{y'_{t+1}}{l^o_{t+1}} - \lambda_{0,t+1} - \frac{\lambda_{4,t+1}}{l^o_{t+1}}\right) = \beta E_t\left(\frac{y'_{t+1}}{l^s_{t+1}}\right) - \beta E_t(\lambda_{0,t+1}) + \lambda_{5,t}$$

Applying (10) to the next period and rewriting gives:

(10'') $$l^o_{t+1} = \frac{l^s_{t+1}}{1-\theta_{t+1}}$$

Inserting into (26) gives:

$$(27)\quad \beta E_t\left(\frac{1-\theta_{t+1}}{l^s_{t+1}}y'_{t+1}-\lambda_{0,t+1}-\lambda_{4,t+1}\frac{(1+\theta_{t+1})}{l^s_{t+1}}\right)=\beta E_t\left(\frac{y'_{t+1}}{l^s_{t+1}}\right)-\beta E_t(\lambda_{0,t+1})+\lambda_{5,t}$$

and finally, because $l^s_{t+1}=l^s_t(1+\pi^s_t)$, we muliply both sides with $-l^s_t$:

$$(28)\quad \beta E_t\left(\frac{\theta_{t+1}}{1+\pi^s_t}y'_{t+1}+l^s_t\lambda_{0,t+1}\right)=l^s_t(-\lambda_{5,t})-\beta E_t\left(\frac{(1-\theta_{t+1})}{(1+\pi^s_t)}\right)\lambda_{4,t}+\beta E_t(l^s_t\lambda_{0,t+1})$$

As $\beta E_t(w+z)=\beta E_t(w)+\beta E_t(z)$, the term $+\beta E_t(l^s_t\lambda_{0,t+1})$ in equation (28) disappears. Using (10a) instead of (10), (28) becomes:

$$(28a)\quad \beta E_t\left(\frac{\vartheta_{t+1}}{(1+\pi^s_t)(1+\vartheta_{t+1})}y'_{t+1}\right)=l^s_t(-\lambda_{5,t})-\beta E_t\left(\frac{(1-\theta_{t+1})}{(1+\pi^s_t)}\right)\lambda_{4,t+1}$$

The left hand side of equation (28) can be interpreted as the expected production-weighted excess return to holding illicit cash rather than legal cash. This expression can either be positive or zero or negative as the right-hand side can turn positive, zero or negative taking into account that $\lambda_{5,t}<\theta$ and $\lambda_{4,t+1}>\theta$ (see above).

- The relative return of illicit cash is the higher, the more binding the condition to hold official cash at a given wage rate in the shadow economy. In a rather realistic world, firms will always hold some legal cash for hiring labor from the official market, otherwise the fiscal authorities will suspect production in the shadow economy. Moreover, even if they could hire all labor required from the (cheaper) parallel market, they will usually hold legal cash for other purposes. This is due to the fact that some cash belongs to the firms's portfolio according to the beliefs of shareholders and tax authorities. The relative return of illicit cash is the lower, the more there is rationing of labor supply in the official labor market. This follows from the endogeneous effect – commented above – that the wage wedge will be narrowed the more labor demand spills over from the official to the parallel labor market.
- The relative return of illicit cash is zero, if the left-hand side of equation (28) is zero. This implies for the right hand side that either $\lambda_{5,t}=\theta$ $\lambda_{4,t+1}=\theta$ or $l^s_t(-\lambda_{5,t})=\beta E_t\left(\frac{(1-\theta_{t+1})}{(1+\pi^s_t)}\right)\lambda_{4,t+1}$. In the *first case*, neither the labor supply constraint on the official market is binding this time nor is the nonnegativity condition ruling. There may be other reasons for holding legal cash, such as legal cash as a long-term store of value. In the *second case*, the positive constraint with regard official money holdings is matched by the rationing of supply in the official labor market.
- When the right hand side of (28) is negative, this implies that the motives for holding a positive amount of legal cash are dominated by a strong "imposed" rationing on the official labor market.

Taking a second-order Taylor approximation of equation (28) gives:

Proposition:

$$(29)\qquad l_t^s(-\lambda_{5,t}) - \beta E_t\left(\frac{(1-\theta_{t+1})}{(1+\pi_t^s)}\right)\lambda_{4,t+1} \cong \beta\left(\frac{E_t\theta_{t+1}}{1+E_t\pi_t^s}\right)y'(E_t i_{t+1}) +$$

$$+\frac{1}{2}\beta\frac{E_t(\theta_{t+1})}{1+E_t\pi_t^s}\left[y'''(E_t i_{t+1})\operatorname{var} i_{t+1} + \frac{2y'(E_t i_{t+1})}{(1+E_t\pi_t^s)^2}\operatorname{var}\pi_t^s\right] +$$

$$+\beta\frac{1}{1+E_t\pi_t^s}y''(E_t i_{t+1})\operatorname{cov}(i_{t+1},\theta_{t+1}) - \beta\frac{E_t\theta_{t+1}}{(1+E_t\pi_t^s)^2}y''(E_t i_{t+1})\operatorname{cov}(i_{t+1},\pi_t^s) +$$

$$-\beta\frac{1}{(1+E_t\pi_t^s)^2}y'(E_t i_{t+1})\operatorname{cov}(\theta_{t+1},\pi_t^s)$$

The development of this second order Taylor series with three variables follows the subsequent rule:

Proof:

$$E[f(i_{t+1},\theta_{t+1},\pi_t^s)] \approx b_0 + b_1\operatorname{var}(i_{t+1}) + b_2\operatorname{var}(\theta_{t+1}) + b_3\operatorname{var}(\pi_t^s)$$
$$+ b_4\operatorname{cov}(i_{t+1},\theta_{t+1}) + b_5\operatorname{cov}(i_{t+1},\pi_t^s) + b_6\operatorname{cov}(\theta_{t+1},\pi_t^s)$$

We achieve the following parameter values:

$$b_0 = f\left[E(i_{t+1}), E(\theta_{t+1}), E(\pi_t^s)\right] = \beta\frac{E(\theta_{t+1})}{1+E(\pi_t^s)}\frac{\partial y[E(i_{t+1})]}{\partial i_{t+1}}$$

$$b_1 = \frac{1}{2}\frac{\partial^2 f\left[E(i_{t+1}), E(\theta_{t+1}), E(\pi_t^s)\right]}{\partial i_{t+1}^2} = \frac{\beta}{2}\frac{E(\theta_{t+1})}{1+E(\pi_t^s)}\frac{\partial^3 y(E(i_{t+1}))}{\partial i_{t+1}^3}$$

$$b_2 = \frac{1}{2}\frac{\partial^2 f\left[E(i_{t+1}), E(\theta_{t+1}), E(\pi_t^s)\right]}{\partial\theta_{t+1}^2} = 0$$

$$b_3 = \frac{1}{2}\frac{\partial^2 f\left[E(i_{t+1}), E(\theta_{t+1}), E(\pi_t^s)\right]}{\partial\pi_t^{s2}} = 2\frac{\beta}{2}\frac{E(\theta_{t+1})}{\left[1+E(\pi_t^s)\right]^3}\frac{\partial y[E(i_{t+1})]}{\partial i_{t+1}}$$

$$b_4 = \frac{\partial^2 f\left[E(i_{t+1}), E(\theta_{t+1}), E(\pi_t^s)\right]}{\partial i_{t+1}\partial\theta_{t+1}} = -\frac{\beta}{2}\frac{1}{1+E(\pi_t^s)}\frac{\partial^2 y[E(i_{t+1})]}{\partial i_{t+1}^2}$$

$$b_5 = \frac{\partial^2 f\left[E(i_{t+1}), E(\theta_{t+1}), E(\pi_t^s)\right]}{\partial i_{t+1}\partial\pi_t^s} = -\beta\frac{E(\theta_{t+1})}{\left[1+E(\pi_t^s)\right]^2}\frac{\partial^2 y[E(i_{t+1})]}{\partial i_{t+1}^2}$$

$$b_6 = \frac{\partial^2 f\left[E(i_{t+1}), E(\theta_{t+1}), E(\pi_t^s)\right]}{\partial\theta_{t+1}\partial\pi_t^s} = \frac{\beta}{2}\frac{1}{\left[1+E(\pi_t^s)\right]^2}\frac{\partial y[E(i_{t+1})]}{\partial i_{t+1}}$$

Q.E.D.

REFERENCES

BARR D.G. and CUTHBERTSON K., "Company Sector Liquid Asset Holdings: A Systems Approach", *Journal of Money, Credit, and Banking*, February 1992, *24*(1), pp. 83-97.

BAUMOL W.J., "The Transactions Demand for Cash: An Inventory Theoretic Approach", *Quarterly Journal of Economics*, November 1952, *66*, pp. 545-56.

BHATTACHARYYA D. K., "An Econometric Method of Estimating the Hidden Economy. United Kingdom (1960-1984): Estimates and Tests", *The Economic Journal*, May 1990, *100*, pp. 703-17.

BOHL M.T. and SELL F.L., "Demand for Cash Balances in Germany. Theoretical Underpinnings and Empirical Evidence", *Applied Economics*, August 1998, *30*(8), pp. 1017-26.

BRÜMMERHOFF D., "Die gesamtwirtschaftliche Finanzierungsrechnung", in W. Ehrlicher and D. Simmert, eds., *Der volkswirtschaftliche Sparprozeß, Beihefte zu Kredit und Kapital*, 1985, Heft 9, pp. 89-111.

DEUTSCHE BUNDESBANK, "Ergebnisse der gesamtwirtschaftlichen Finanzierungsrechnung für Westdeutschland 1960 bis 1992", *Statistische Sonderveröffentlichungen*, Nr. 4, Frankfurt/Main, 1994.

——, "Ergebnisse der gesamtwirtschaftlichen Finanzierungsrechnung für Deutschland 1990 bis 1998", *Statistische Sonderveröffentlichungen*, Nr. 4, Frankfurt/Main, 1999.

FALK M. and FUNKE N., "The Stability of Money Demand in Germany and in the EMS: Impact of German Unification", *Weltwirtschaftliches Archiv*, September 1995, *131*, pp. 470-88.

FASE M.M.G. and WINDER C.C.A., "The Demand for Money in the Netherlands Revisited", *De Economist*, 1990, *138*(3), pp. 276-301.

HANSEN G. and Kim H.R., "The Stability of German Money Demand: Tests of the Cointegration Relation", *Weltwirtschaftliches Archiv*, June 1995, *131*(2), pp. 276-301.

HINZE J., "Relation zwischen Bruttoarbeitskosten und Nettolöhnen", HWWA Report No. 202, Hamburg Institute of International Economics, Hamburg, 2000.

JOHANSEN S., "Statistical Analysis of Cointegration Vectors", *Journal of Economic Dynamics and Control*, 1988, *12*, pp. 231-54.

——, "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models", *Econometrica*, 1991, *59*, pp. 1551-80.

KARMANN A., "Monetäre Schätzansätze zur Erfassung der Schattenwirtschaft. Ein Vergleich verschiedener Meßmethoden", *Kredit und Kapital*, 1986, *19*(2), pp. 233-47.

——, "Schattenwirtschaft und ihre Ursachen. Eine empirische Analyse zu Schwarzwirtschaft und Selbstversorgung in der Bundesrepublik Deutschland", *Zeitschrift für Wirtschafts- und Sozialwissenschaft*, 1990, *110*(2), pp. 185-206.

KEYNES J.M., *The General Theory of Employment, Interest and Money*, New York: Macmillan, 1936.

KREPS D.M., *Mikroökonomische Theorie*, Landsberg/Lech: Verlag Moderne Industrie, 1994.

KWIATKOWSKI D., PHILLIPS P.C.B., SCHMIDT P., and SHIN Y., "Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root", *Journal of Econometrics*, 1992, *54*, pp. 159-78.

LANE T.D. (1990a), "The Demand for Money in Poland: Theory and Evidence", Working Paper, IMF Research Department, April 1990.

—— (1990b), "Costly Portfolio Adjustment and the Short-Run Demand for Money", *Economic Inquiry*, July 1990, *28*(3), pp. 466-87.

——, "Household Demand for Money in Poland", *IMF Staff Papers*, December 1992, *39*(4), pp. 825-54.

LEWIS W.A., "Economic Development with Unlimited Supplies of Labour", *The Manchester School of Economic and Social Studies*, 1954, *22*, pp. 139-91.

MELTZER A.H., "The Demand for Money: A Cross-Section Study of Business Firms", *The Quarterly Journal of Economics*, 1963, *77*, pp. 405-22.

MILLER M.H. and ORR D., "A Model of the Demand for Money by Firms", *Quarterly Journal of Economics*, August 1966, *80*, pp. 413-35.

MIZEN P., "Modeling the Demand for Money in the Industrial and Commercial Companies Sector in the United Kingdom", *Journal of Policy Modeling*, 1996, *18*(4), pp. 445-67.

MULLIGAN C.B., "Scale Economies, the Value of Time, and the Demand for Money: Longitudinal Evidence from Firms", *Journal of Political Economy*, October 1997, *105*(5), pp. 1061-79.

NICHOLSON W., *Microeconomic Theory. Basic Principles and Extensions*, 5th ed., Fort Worth: Dryden Press, 1992.

OSTERWALD-LENUM M., "A Note with Quantiles of the Asymptotic Distribution of the Maximum Likelihood Cointegration Rank Test Statistics", *Oxford Bulletin of Economics and Statistics*, 1992, *54*, pp. 461-71.

SACHVERSTÄNDIGENRAT, *Jahresgutachten 2001/02 des Sachverständigenrates zur Begutachtung der gesamtwirtschaftlichen Entwicklung*, Stuttgart: Metzler-Poeschel, 2001.

SCHNEIDER F., "Schwarzarbeit legt um 7,5% zu", *Handelsblatt*, 13.12.1999, p. 6.

SCHWERT G.W., "Tests for Unit Roots: A Monte Carlo Investigation", *Journal of Business and Economic Statistics*, 1987, *7*, pp. 147-59.

SEITZ F., "Der DM-Umlauf im Ausland", Diskussionspapier 1, Volkswirtschaftliche Forschungsgruppe der Deutschen Bundesbank, Frankfurt/Main, 1995.

SELL F.L., "Zu den langfristigen Bestimmungsgründen der Umlaufgeschwindigkeit des Geldes: Ein sektoraler Ansatz", *Kredit und Kapital*, 1997, *30*, pp. 325-47.

TAYLOR M.P., "The Hyperinflation Model of Money Demand Revisited", *Journal of Money, Credit, and Banking*, August 1991, *23*(3), pp. 326-51.

TOBIN J., "The Interest-Elasticity of Transactions Demand for Money", *The Review of Economics and Statistics*, 1956, *38*, pp. 241-47.

VIREN M., "Demand for Cash by Business Firms", Working Paper, Bank of Finland, 1996.

VOGEL R.C. and MADDALA G.S., "Cross-Section Estimates of Liquid Asset Demand by Manufacturing Corporations", *Journal of Finance*, 1967, *22*, pp. 557-75.

WAHLEN E.L., "A Cross-Section Study of Business Demand for Cash", *Journal of Finance*, 1965, *20*, pp. 423-43.

WOLTERS J., TERÄSVIRTA T., and LÜTKEPOHL H., "Modeling the Demand for M3 in the Unified Germany", *The Review of Economics and Statistics*, August 1998, *80*(3), pp. 399-409.

ABSTRACT

Based on a cash-in-advance approach, this paper investigates theoretically the determinants of money holdings of firms under the conditions of a highly regulated labor market and analyses empirically the demand for money of German businesses during the period 1960-1998. As a result of our theoretical analysis the demand for cash balances by firms for shadow market activities depends among other things positively on the expected wage wedge. The empirical results show that the coefficient of the wage wedge has a positive sign in the long-run cointegrating relationship and is statistically significantly positive in the short-run dynamics of the error correction model.

JEL classification: *E41, C22, J30*
Keywords: *money demand by firms, wage wedge, cash-in-advance model, cointegration, error-correction*

RISEC, Volume 50 (2003), No. 4, 475-484

QUALITY OF ASSETS, BANK PROFITABILITY AND CAPITAL POSITION

by
ATHANASIOS G. NOULAS* and JOHN A. PAPANASTASIOU*

1. *Introduction*

Banks have to make decisions on the amount of capital they hold for a number of reasons. First, by law, banks must have a minimum amount of capital in order to satisfy the regulatory authorities. Second, the amount of capital plays the role of a cushion against bank failure. The larger the amount of capital, the lower the risk that the bank will become insolvent. Third, the amount of capital affects the return on equity. Given the return on assets, the lower the amount of capital, the higher the return on equity. Thus, the management of capital is important since capital is one of the factors that provide confidence to a bank and in general to the banking system. Of course how much capital a bank needs in order to ensure confidence to the public, to investors, to creditors and to regulators is another question. As with many other questions the correct answer is: "It depends on...". It depends on many factors such as the liquidity position, the profitability, the size, and the quality of the management.

The basic question we try to answer in this paper is this: What is the relationship between bank profitability, quality of assets and capital position? Also, we investigate the causality between quality of assets and profitability for the first time in the Greek banking system.

A number of studies, mainly on U.S data, have examined the behavior of

* Department of Accounting and Finance, University of Macedonia, Thessaloniki, (Greece). E-mail anoulas@uom.gr

We would like to thank the anonymous RISEC referee for the constructive and helpful comments.

banks to capital requirements as well as the relationship between capital-asset ratio and profitability. Kahane (1977), Koehn and Santomero (1980), and Kim and Santomero (1988) have shown that increases in capital standards forced by regulatory authorities positively affect the portfolio risk. On the other hand, studies by Dothan and Williams (1980), and Benston et al. (1986) have indicated that capital and portfolio risk may be negatively correlated. Rochet (1992) has found that capital regulations cannot prevent banks from following a risky behavior if they try to maximize the value of their future profits. However, if the portfolio model is adopted then capital regulations can be effective. Jacques and Nigro (1997) have examined the impact of risk-based capital standards on capital ratios and portfolio risk for 2570 FDIC-insured commercial banks. The study, which covers the first year the banks were required to follow the new rule on capital, has found that the new risk-based capital caused an increase in the capital ratios and a reduction in the portfolio risk. Rime (2001) examined the reaction of banks to regulatory constraints on capital. The study, using data for Swiss banks for the period 1989-1995, has found that regulatory pressures have a positive effect on the capital-asset ratio and no significant effect on the behavior of banks toward risk.

Berger (1995) using U.S banks data for the decade of the 1980s has examined the relationship between capital and profitability and found a strong positive relationship. Moreover, each of these two variables positively Granger-caused the other. Accordingly, an increase in profitability causes an increase in capital and vice versa. For the early 1990s, however, the author does not find a positive Granger-causality from capital to earnings, a result that is attributed by the author to the possibility that the banks had more than the optimal level of capital during this period. The last result is consistent with the bankruptcy hypothesis, which states that when capital is below its optimal level there is a positive causality from capital to earnings and a negative causality when capital exceeds its optimal level.

2. *Methodology and Data*

We separate the banks into two groups based on their capital position. In other words, we argue that there are two distinct groups of banks that follow different policies regarding the amount of capital. Based on the strategy the management follows regarding capital decisions, banks are designated as aggressive and non-aggressive (obviously, they could be also called high and low capital banks). If the capital-asset ratio is smaller than the industry's

median, then a bank is considered to be an aggressive one. If, on the other hand, the capital-asset ratio is larger than the industry's median, the bank is considered to be non-aggressive.

The model. - The model we consider draws from the regression-switching regime with known sample separation literature (see Maddala, 1983, ch. 10). Specifically, the assets quality variable (Q), is described by a vector x of a priori assumed weakly exogenous variables in two different ways depending upon whether or not a bank follows a non-aggressive policy. Letting the subscript NA denote the first regime under which the variable Q operates and A the complimentary regime, then the function of the assets quality of bank i may be characterized as:

$$Q_i = x'_{i,NA}\beta_{NA} + \varepsilon_{i,NA} \qquad \text{if bank } i \text{ follows a non-aggressive policy}$$

$$Q_i = x'_{i,A}\beta_A + \varepsilon_{i,A} \qquad \text{otherwise}$$

where x and β are appropriately and conformably dimensioned vectors of variables and parameters corresponding to the structural description of assets quality under the non-aggressive and aggressive policy regimes. This is a reasonably familiar switching regime model with known sample separation in the sense that whether a bank's capital decision is aggressive or not is available information. The wrinkle here is that the capital position is not independent of the quality of assets decision. To get consistent estimates we utilize a two-stage estimation procedure. The structure of the two-stage technique is characterized as:

$$(1) \qquad Q_i = x'_{i,NA}\,\beta_{NA} + \varepsilon_{i,NA} \qquad \text{if } \gamma' y_i \geq \varepsilon_i$$

$$(2) \qquad Q_i = x'_{i,A}\,\beta_A + \varepsilon_{i,A} \qquad \text{otherwise}$$

where $\gamma' y_i$ is a stochastic criterion function of variables y_i, and parameters γ determine whether a bank i is aggressive or not, so that an indicator function I_i is posited such that:

$$I_i = 1 \qquad \text{if } \gamma' y_i \geq e_i \text{ and bank } i \text{ is non aggressive}$$

$$I_i = 0 \qquad \text{otherwise}$$

where e_i is the median of the capital adequacy index (K) and γ can be estimated up to a scale factor by probit methods. Furthermore, it is assumed

$\varepsilon_{i,NA}$, $\varepsilon_{i,A}$, and e_i to be trivariate normally distributed with zero mean and covariance matrix of the form:

$$\begin{bmatrix} \sigma^2_{NA} & \sigma^2_{NA,A} & \sigma^2_{NA,e} \\ \sigma^2_{A,NA} & \sigma^2_{A} & \sigma^2_{A,e} \\ \sigma^2_{e,NA} & \sigma^2_{e,A} & 1 \end{bmatrix}$$

It follows that:

$$E(\varepsilon_{i,NA} \mid e_i \leq \gamma' y_i) = E(\sigma_{NA,e} e_i \mid e_i \leq \gamma' y_i) = -\sigma_{NA,e} \frac{\varphi(\gamma' y_i)}{\Phi(\gamma' y_i)} = -\sigma_{NA,e} w_{i,NA}$$

and in a similar way

$$E(\varepsilon_{i,A} \mid e_i \geq \gamma' y_i) = E(\sigma_{A,e} e_i \mid e_i \geq \gamma' y_i) = \sigma_{A,e} \frac{\varphi(\gamma' y_i)}{1 - \Phi(\gamma' y_i)} = \sigma_{A,e} w_{i,A}$$

where $\varphi(\ .\ )$ and $\Phi(\ .\ )$ are respectively the standard normal density function and its cumulant evaluated at its argument (see Maddala, 1983, p. 224). Therefore, equations (1) and (2) may be written as:

(3) $$Q_i = x'_{i,NA} \beta_{NA} - \sigma_{NA,e} w_{i,NA} + u_{i,NA} \quad \text{for } I_i = 1$$

(4) $$Q_i = x'_{i,A} \beta_A - \sigma_{A,e} w_{i,A} + u_{i,A} \quad \text{otherwise}$$

where the residuals u are implicitly related to the residuals ε in equations (1) and (2). Thus, the two-stage method utilizes probit analysis in the first stage and a least squares procedure in the second stage. Specifically, in the first stage probit estimates of γ permit to establish values for w_{NA} and w_A for each i via evaluation of the functions $\varphi(\ .\ )$ and $\Phi(\ .\ )$. In the second stage, the equations (3) and (4) are estimated by weighted least squares since the residuals u are heteroskedastic. This procedure provides consistent estimates of the parameters. Thus, (3) and (4) are familiar regression equations with heteroskedastic errors of a known form. The conditional variances of these errors may be shown to be:

(5) $$\sigma^2(u_{i,NA} \mid I_i = 1) = \sigma^2_{NA} - \sigma^2_{NA,e} w_{i,NA} (\gamma' y_i + w_{i,NA})$$

$$(6) \qquad \sigma^2(u_{i,A} \mid I_i = 0) = \sigma_A^2 + \sigma_{Ae}^2 w_{i,A}(\gamma' y_i - w_{i,A})$$

Thus, equations (5) and (6) can be used to obtain the appropriate heteroskedasticity adjustments since all that is required in addition to the first round estimates are estimators for the variances which are given by[1]:

$$(7) \qquad \hat{\sigma}_{NA}^2 = T_{NA}^{-1} \sum_{i=1}^{T_{NA}} \left[\hat{u}_{i,NA}^2 + \hat{\sigma}_{NA,e}^2 (\hat{\gamma}' y_i) \hat{w}_{i,NA} + \hat{\sigma}_{NA,e}^2 \hat{w}_{i,NA}^2 \right]$$

$$(8) \qquad \hat{\sigma}_A^2 = T_A^{-1} \sum_{i=1}^{T_A} \left[\hat{u}_{i,A}^2 - \hat{\sigma}_{A,e}^2 (\hat{\gamma}' y_i) \hat{w}_{i,A} + \hat{\sigma}_{A,e}^2 \hat{w}_{i,A}^2 \right]$$

where T_{NA} is the number of observations for which $I_i = 1$ and T_A is the number of observations for which $I_i = 0$.

It should be pointed out that since the variable Q (quality of assets) does not appear in the vector y, if $\sigma_{NA,e}$ and $\sigma_{A,e}$ are both equal to zero then the switching process may be thought of as exogenous. Expressed in another way the variable Q does not "cause" y and thus y may be constructed as independent of Q.

Consequently if the estimated coefficients $\sigma_{NA,e}$ and $\sigma_{A,e}$ of equations (3) and (4) are statistically not significant then there is no correlation between ε and e which implies that Q does not "cause" y.

The vector y contains two variables: profitability as measured by the return on assets (*ROA*) and the bank size as measured by total assets (A). We hypothesize here that capital decision depends on the profitability and the size of the bank. We expect banks with lower profitability (*ROA*) to be more aggressive on their capital position. The reason is that the smaller capital position means higher leverage (equity multiplier), which boosts the return on equity. With respect to the size, it could be argued that the larger the bank the more aggressive it is. This is due to the fact that the larger bank is better diversified and equipped to deal with serious or difficult situations since it has the ability to attract high quality labor.

The variables that are included in the vector x are the following:

The ratio of bonds to total assets (B). It is well known that banks invest in bonds for the purposes of liquidity and income. The risk involved with bonds

---

[1] These formulas drawn from Lee and Trost (1978) guarantee that the estimates $\hat{\sigma}_{NA}^2$ and $\hat{\sigma}_A^2$ are always positive. For demonstration of formulas (7) and (8) see Appendix.

is the interest rate risk. Increases in interest rates cause the price of bonds to fall especially for long-term bonds. The bonds held by the Greek banks are basically short-term government bonds and therefore the interest rate risk is low. In addition, the period 1993-1998 is a period of gradual reduction in the Greek interest rates for the effort of the country to satisfy one of the criteria to become a member of the Economic and Monetary Union. This means that the expected capital gains made the bonds a better investment than loans. Therefore, we expect a positive relation between bonds and quality of assets.

Return on Assets (ROA). On a theoretical basis it could be argued that banks with a higher capital position may decide to relax the standards on the loan policy and pay more attention to the quantity than to the quality as profitability increases and things are getting better. On the other hand, banks with a lower or poor capital position being therefore under the scrutiny of the regulatory authorities and the market, may have to pay more attention to the quality of assets. Of course the counterargument could be made here that banks with a poor capital position having problems with operating revenue, may "bet" the bank by deciding to give out more risky loans and higher expected return. If the gamble pays out, it will save the bank from big problems, if it doesn't, the bank will not loose much since its situation was not good anyhow. Accordingly, the effect of profitability on the quality of assets is an open question to be determined by empirical evidence.

Ownership status (D). This is a dummy variable that takes the value of one if the bank is controlled by the State, and zero if the bank is privately owned.

The data. - We use year-end data for sixteen Greek banks for the years 1993-1998 taken from the balance sheets. The definition and measurement of the variables are as follows. The size (A) is measured by the logarithm of total assets; the capital position (K) is measured by the ratio of capital to total assets.[2] The quality of assets (Q) is measured by the ratio of provision for loan losses to total loans.[3] D is a dummy variable that captures the ownership status. For example, $D = 1$ if the bank is controlled by the State, zero otherwise. Investment in bonds (B) is measured by the ratio of bonds to total assets, and profitability (ROA) is measured by the return on assets.

[2] Capital includes only equity capital which is the sum of common stock, surplus and retained earnings.

[3] Loan loss provision is the only item that is provided in the data.

3. *Empirical Results*

Profitability, size and capital position

Probit estimates of the capital decision function:

$$I = \underset{(3.07)}{41.28}\, ROA - \underset{(2.49)}{0.04}\, A$$

$R^2 = 0.13$, LR (zero slopes) = 10.17 [0.001]

The coefficient of *ROA* is significant and positive indicating that banks with low profitability are more likely to be aggressive in their capital decision. As expected, low-profitability banks have an incentive to have a higher degree of leverage so that they can magnify the return on equity. The coefficient of *A* is significant and negative implying that large banks are more likely to be aggressive in their capital decisions.

Quality of assets equation: Non-Aggressive Banks

$$Q = \underset{(2.57)}{-0.0195}\, B + \underset{(4.76)}{0.268}\, ROA + \underset{(7.88)}{0.0124}\, D + \underset{(3.08)}{0.0067}\, W_1$$

$R^2 = 0.63$, F = 21.56 [0.000], HET = 12.15 [0.000], NORM = 1.86 [0.394], RESET = 5.73 [0.022]

where HET, NORM and RESET are the tests for heteroscedasticity, normality and functional misspecification respectively.

Quality of assets equation: Aggressive Banks

$$Q = \underset{(2.33)}{-0.0161}\, B - \underset{(2.31)}{0.35}\, ROA + \underset{(7.09)}{0.0125}\, D + \underset{(4.1)}{0.016}\, W_2$$

$R^2 = 0.52$, F = 13.91 [0.000], HET = 5.44 [0.020], NORM = 1.37 [0.504], RESET = 3.34 [0.076]

The coefficient of *B* (ratio of bonds to total assets) is significant and negative for both groups of banks indicating that *B* affects the quality of loans. As the investment in bonds increases the quality of assets increases (the quality index declines). Why are higher levels of bonds associated with higher quality of assets? An increase in government bonds, ceteris paribus, means that loans as portion of total deposits decline and with it the amount for loan loss provision.

The coefficient of *ROA* is positive and significant for non-aggressive banks indicating that higher profitability is associated with lower quality. For aggressive banks the opposite is true. Higher profitability is associated with higher quality. At first glance this result might seem contradictory for the two sets of banks. However, a careful examination leads to the following explanation. Non-aggressive banks are banks with above the average capital to asset ratio and better capital position means that the risk of insolvency and bankruptcy is smaller than the risk that aggressive banks face. This higher capital position allows them to invest in loans with higher risk. The opposite is true for banks with low capital position. An increase in profitability is an opportunity for them to reconstruct their portfolio of loans making it healthier by paying more attention to the quality. The dummy variable capturing the status of ownership is the same for both sets of banks.

The coefficients of W_1 and W_2 are highly significant implying the existence of causality effects from quality to profitability. That is, it is the quality of assets, which "causes" profitability in the Greek banks.

4. *Conclusion*

Separating banks into two groups based on their capital position, we found that the effect of profitability on quality is different for the two groups. For non-aggressive banks the increase in profitability is associated with lower quality while the opposite is true for aggressive banks.

The results showed that the relationship between quality and profitability is a two-way street. Quality "causes" profitability but profitability then affects quality. Since it seems to exist a simultaneity effect between quality and *ROA*, and since *ROA* has been used as an explanatory variable of quality, the implication here is that a system of equations should be used in empirical studies to take into account the simultaneity effect between quality and profitability.

APPENDIX

From Lee-Trost (1978) it is known that:

$$E(\varepsilon_{i,NA} \mid I_i = 1) = -\sigma_{NA,e}\, w_{i,NA} \tag{9}$$

$$E(\varepsilon_{i,A} \mid I_i = 0) = \sigma_{A,e}\, w_{i,A} \tag{10}$$

$$E(\varepsilon^2_{i,NA} \mid I_i = 1) = \sigma^2_{NA} - \sigma_{NA,e}\,(\gamma' y_i)\, w_{i,NA} \tag{11}$$

$$E(\varepsilon^2_{i,A} \mid I_i = 0) = \sigma^2_{A} + \sigma_{A,e}\,(\gamma' y_i)\, w_{i,A} \tag{12}$$

From equations (3) and (4) we have:

$$u_{i,NA} = \varepsilon_{i,NA} + \sigma_{NA,e}\, w_{i,NA} \quad \text{for } I_i = 1 \tag{13}$$

$$u_{i,A} = \varepsilon_{i,A} - \sigma_{A,e}\, w_{i,NA} \quad \text{for } I_i = 0 \tag{14}$$

Raising equations (13) and (14) into square power and taking the expectations, given equations (9) and (10) we obtain:

$$E(u^2_{i,NA} \mid I_i = 1) = E(\varepsilon^2_{i,NA} \mid I_i = 1) - \sigma^2_{NA,e}\, w^2_{i,NA} \tag{15}$$

$$E(u^2_{i,A} \mid I_i = 0) = E(\varepsilon^2_{i,A} \mid I_i = 0) - \sigma^2_{A,e}\, w^2_{i,A} \tag{16}$$

Substituting equations (11) and (12) into equations (15) and (16) respectively, we take equations (5) and (6). Solving equations (5) and (6) with respect to σ^2_{NA} and σ^2_{A} we derive estimates for the variances (equations (7) and (8)).

REFERENCES

BENSTON G. J., EISENBEIS R. A., HORVITZ P. M., KANE E. J., and KAUFMAN, G. G., *Perspectives on Safe and Sound Banking*, Cambridge, MA: MIT Press, 1986.

BERGER A.N., "The Relationship Between Capital and Earnings in Banking", *Journal of Money, Credit, and Banking*, 1995, *27*(2), pp. 432-56.

DOTHAN U. and WILLIAMS J., "Banks, Bankruptcy, and Public Regulation", *Journal of Banking and Finance*, 1980, *4*(1), pp. 51-60.

JACQUES K. and NIGRO P., "Risk-Based Capital, Portfolio Risk, and Bank Capital: A Simultaneous Equations Approach", *Journal of Economics and Business*, 1997, *49*, pp. 533-47.

KAHANE Y., "Capital Adequacy and the Regulation of Financial Intermediaries", *Journal of Banking and Finance*, 1977, *1*(2), pp. 207-17.

KIM D. and SANTOMERO A.M., "Risk in Banking and Capital Regulation", *Journal of Finance* 1988, *43*, pp. 1219-33.

KOEHN M. and SANTOMERO A.M., "Regulation of Bank Capital and Portfolio Risk", *Journal of Finance*, 1980, *35*(5), pp. 1235-44.

LEE L.F. and TROST R.P., "Estimation of Some Limited Dependent Variable Models with Applications to Housing Demand", *Journal of Econometrics*, 1978, *8*, pp. 357-82.

MADDALA G.S., *Limited-Dependent and Quantitative Variables in Econometrics*, Cambridge: Cambridge University Press, 1983.

RIME B., "Capital Requirements and Bank Behaviour: Empirical Evidence for Switzerland", *Journal of Banking and Finance*, 2001, *25*, pp. 789-805.

ROCHET J.C., "Capital Requirements and the Behavior of Commercial Banks", *European Economic Review*, 1992, *36*, pp. 1137-78.

ABSTRACT

The paper examines the relationship between bank profitability, quality of assets and capital position using a switching regression model with known sample separation. It also investigates the causality between quality of assets and profitability of banks. Banks are separated into aggressive and non-aggressive based on their capital position. We find that the effect of profitability on quality is different for the two groups. For non-aggressive banks the increase in profitability has a negative impact on the quality while for aggressive banks the relationship is positive. Quality causes profitability but profitability then affects quality.

JEL classification: *G21, G28*
Keywords: *bank profitability, quality of assets, capital position, causality, switching regime model.*

RISEC, Volume 50 (2003), No. 4, 485-502

OPERATIONAL EFFICIENCY IN THE GREEK BANKING INDUSTRY: EVIDENCE FROM THE STOCHASTIC FRONTIER APPROACH AND ERROR DECOMPOSITIONS

by
NICHOLAS APERGIS* and ANTHONY REZITIS**

1. *Introduction*

Extensive changes are taking place in the financial industry around the world. The United States is moving toward nationwide banking, while in Asia the regulations governing the financial industry are changing in an effort to make it more efficient. The European Committee issued a White Paper in 1985 through which 300 directives were provided to achieve a unified market for goods and financial services. This financial integration is expected to lead to the disappearance of all capital controls. Financial institutions will be allowed to operate throughout the European Union (EU) under a single banking license. At the same time, the deregulatory activities will intensify the degree of competition across the EU countries. In such a multinational competitive environment, bank efficiency becomes an important issue, as inefficient banks are expected not to survive.

---

* Department of Economics, University of Macedonia, Thessaloniki (Greece).

** Department of Farm Organization and Management, University of Ioannina, Ioannina (Greece). E-mail: arezitis@cc.uoi.gr

Paper presented at the 49th International Atlantic Economic Conference, Munich (Germany).

The authors would like to express their gratitude for the helpful comments of Prof. Theodore Palivos as well as the participants of the session: Financial and Capital Markets in Greece, on an earlier draft of this paper.

The goal of this paper is to measure X-efficiency of the Greek banking sector. The main characteristic of the Greek banking system during the 1980s was the operation of the system under the complete dominance of the monetary authorities, which resulted in the absence of a competitive banking environment and the limited role the banks were called to play in the economy's growth process. This sector, however, has been characterized by dramatic and substantial changes, which have occurred since 1988. Such changes involved: deregulation of interest rates, abolishment of credit rationing, higher capital mobility, abolishment of the requirement by banking institutions to invest a significant portion of their reserves to public bonds, and free entry of banking institutions from abroad, mainly from the EU economies. Nevertheless, these changes have not found (completely) their way to the banking sector, even today. On the other hand, if the deregulation creates more 'bad' loans in the banking industry, then these loans have an adverse impact on the efficiency of the industry since they are associated with higher monitoring expenses and higher management risks (Fukuyama, 1995; Berger and Humphrey, 1997).

This is the first study, that uses a stochastic frontier approach (SFA) to examine X-efficiency of the Greek banking industry. The term X-efficiency is extensively used in the banking literature (Berger et al., 1993; Berger and Humphrey, 1997) and refers to both allocative and technical efficiencies of individual banks, as opposed to scale and scope efficiencies. More specifically, X-efficiency implies that banks optimize the input mix in order to avoid both excessive levels of input usage (technical X-inefficiency) as well as non-optimal usage of input proportions (allocative X-inefficiency). On the other hand, scale and scope efficiencies require that banks optimize the output mix in order to exploit all scale and scope economies. Empirical evidence from US commercial banks (Berger et al., 1993; Berger and Humphrey, 1997) suggests that X-inefficiency accounts for about 20% or more of banking costs as opposed to scale and scope inefficiencies, which count for less than 5% of costs. These findings indicate that X-inefficiency appears to be a much more important source of deviations from the efficient frontier than scale and scope inefficiencies. Thus, since this work is the first in studying efficiency in the Greek banking, it appears more reasonable to examine X-efficiency rather than scale and scope efficiencies.

The present study uses a SFA, introduced by Aigner et al. (1977), which combines both technical and allocative X-efficiencies in the cost relationship. In measuring X-efficiency, this approach assumes two types of error in the production process. The first type accounts for systematic deviations from the stochastic frontier due to inefficiencies in the production process. The second

type accounts for random variations in the costs, due to factors outside the control of the bank manager. An additional goal of the present study is to examine the relative contribution of the factors that affect X-efficiency of the Greek banking industry. This can be achieved by decomposing the bank-specific X-inefficiency measures derived from the estimation of the stochastic cost frontier into various bank characteristics.

Bank efficiency is important because, the more efficient the banks are, the better the functioning of the whole economic system. It is beyond any doubt that the banking system plays a pivotal role in determining how well the rest of the economy performs and in particular the growth prospects of an economy. A more efficient financial system promotes growth for two main reasons. In the first place, a larger financial penetration guarantees that funds are allocated to those projects with higher rates of return. A more developed financial system helps the process of gathering the information that connects lenders and borrowers and hence reduces the informational asymmetries that arise in capital markets (Greenwod and Jovanovic, 1990; Saint Paul, 1992; Pagano, 1993). At the same time, the amount of funds allocated to productive investment will be larger if the system is more efficient, since the intermediation margin retained by the financial institutions will be smaller. This margin in turn may reflect inefficiencies and market power of the banks as well as government regulations, such as reserve requirements or taxes (Roubini and Sala-i-Martin, 1992). Furthermore, in a competitive environment, inefficient banks cannot survive for a long time. Thus, bank management must undertake additional actions that will lead toward efficient production.

The remaining of the paper is organized as follows: Section 2 provides a brief presentation of the international research attempts made in the literature to measure banking inefficiency, while Section 3 provides a brief description of the methodology followed in this paper. Section 4 presents the empirical results, and Section 5 concludes.

2. *Literature Review*

Studies concerned with banking efficiency estimate a 'best-practice cost frontier', which represents the cost frontier of efficient banks. The degree of inefficiency of banks in the sample is measured by the deviation from the 'best-practice cost frontier'. Five common approaches have been used in evaluating the efficiency of financial institutions. These methods include the data envelopment analysis (DEA), the free disposal hull (FDH), the thick

frontier approach (TFA), the distribution-free approach (DFA), and the stochastic frontier approach (SFA). The first two approaches are nonparametric, while the other ones are parametric.

Nonparametric frontiers employ the principles of linear programming and assume that all deviations from the constructed frontier represent inefficiencies, because nonparametric approaches do not generally account for random errors. Some nonparametric studies that have used DEA include Favero and Papi (1995) for Italian banks, Taylor et al. (1997) for Mexican banks, Fukuyama (1995) for Japanese banks, Berg et al. (1991) for Norwegian banks, Zaim (1995) for Turkish banks, and Miller and Noulas (1996) for US banks. A few studies that employ FDH include Tulkens (1993) for Belgian banks and DeBorger et al. (1995) for US banks.

Parametric frontiers specify a functional form for the 'best-practice frontier' and they account for random errors. The parametric methods differ between each other in terms of the way they separate inefficiencies from random errors. The thick frontier approach (TFA) divides the banks in the sample into quartiles based on average costs. It assumes that differences in the predicted cost functions between the lowest and highest average cost quartiles represent inefficiency, while deviations from predicted costs within each average cost quartile represent random errors. Studies that use TFA include Lang and Welzel (1995) for German banks, Berg and Kim (1994) for Norwegian banks, Lozano (1995) for Spanish banks and Bauer et al. (1993) for US banks. The distribution free approach (DFA) assumes that efficiency differences are stable over time, whereas random error tends to average out to zero over time. Studies that use DFA include Dietsch (1994) for French banking, and Lang and Welzel (1996) for German banks.

The stochastic frontier approach (SFA), which is used in this study, specifies a functional form for the cost relationship among inputs and outputs and allows for the error term to be composed of two terms. The first term captures measurement errors and other factors beyond the control of the firm, while the second term measures inefficiency. The majority of the literature has investigated the efficiency of the U.S. banking industry. A few studies have examined the efficiency of banks outside the U.S. and a handful of them utilise the stochastic frontier approach (SFA). In particular, the paper by Berger and Humphrey (1997) surveys 130 studies that use various frontier efficiency methods to examine efficiency of financial institutions in 21 countries. This work shows that a number of about 19 studies apply the stochastic frontier approach (SFA), to examine efficiency of the banking industry. The majority of this literature, a number of about 13 studies, investigates the efficiency of the U.S. banking industry. Some of these

studies include Ferrier and Lovell (1990), Allen and Rai (1993), Bauer et al. (1993), Chang et al. (1993), Kaparakis et al. (1994), Kwan and Eisenbeis (1994), Eisenbeis et al. (1996), and Mester (1996, 1997). A few studies, about 6, examine the efficiency of banks outside the U.S. Such studies include Chaffai (1993, 1997) for Tunisian banks, Resti (1995) for Italian banks, and Maudos (1996a, 1996b) for Spanish banks.

A parametric approach, i.e. the SFA, is used in this study instead of a nonparametric approach, i.e. DEA or FDH, because the latter does not take into account random errors which may temporarily give some banking units high or low costs, profits, outputs or inputs. In addition, nonparametric approaches use only data on inputs and outputs and do not take direct account of input prices, thus, they do not measure allocative X-inefficiency (Berger, 1993). Finally, the SFA is preferred over other parametric approaches, i.e. TFA or DFA, because the latter require a large number of cross sectional units while in the present study only six banking units are available.

3. *Methodological Issues*

3.1. *The SFA approach.* - In this paper, a method, proposed by Aigner et al. (1977), is used to estimate a stochastic cost frontier for the Greek banking industry. The basic model assumes that a bank's observed cost deviates from the cost frontier because of a random noise, v_i, and an inefficiency component, u_i. That is, for N banks in the sample the stochastic cost frontier function is

$$\ln C_i = \ln C(y_i, w_i) + \varepsilon_i \qquad i = 1, \ldots, N \tag{1}$$

where C_i is the observed cost of bank i, y_i is the vector of outputs for bank i, w_i is the vector of input prices for bank i, and ε_i is a disturbance term consisting of two independent elements as follows

$$\varepsilon_i = v_i + u_i \tag{2}$$

The symmetric component v_i captures the effect of random factors outside the bank's control, measurement errors in the dependent variable, and any other statistical 'noise'. It is also assumed to be independently and identically distributed as $N(0, \sigma_v^2)$, i.e. v_i follows a normal distribution. The one-sided component $u_i \geq 0$ reflects cost inefficiency relative to the stochastic frontier.

Thus, $u_i = 0$ for a bank whose cost lies on the frontier, and $u_i > 0$ for one whose cost is above the frontier. It is assumed that u_i is identically and independently distributed as $|N(0, \sigma_u^2)|$, i.e. u_i follows a half-normal distribution. With these distributional assumptions, the log-likelihood function of the stochastic frontier model is

$$\ln L = \frac{N}{2}\ln\frac{2}{\pi} - N\ln\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{N}\varepsilon_i^2 + \sum_{i=1}^{N}\ln\left[\Phi\left(\frac{\varepsilon_i\lambda}{\sigma}\right)\right] \tag{3}$$

where N is the number of banks, $\sigma^2 = \sigma_u^2 + \sigma_v^2$, $\lambda = \sigma_u/\sigma_v$, and $\Phi(.)$ is the standard normal cumulative distribution function evaluated at $\varepsilon_i\lambda/\sigma$. The model can be estimated using maximum likelihood techniques. Jondrow et al. (1982) demonstrate that inefficiency measures can be calculated using the residuals of the estimated model.

Specifically, bank-specific inefficiency measures are given by

$$E(u_i \mid \varepsilon) = \left[\frac{\sigma\lambda}{1+\lambda^2}\right]\left[\frac{\varphi\left(\frac{\varepsilon_i\lambda}{\sigma}\right)}{\Phi\left(\frac{\varepsilon_i\lambda}{\sigma}\right)} + \frac{\varepsilon_i\lambda}{\sigma}\right] \tag{4}$$

where $\phi(.)$ is the standard normal density function evaluated at $\varepsilon_i\lambda/\sigma$.

The coefficient $\lambda = \sigma_u/\sigma_v$ shows the relative variation of the standard error of u_i to the standard error of v_i. Values of λ close to zero indicate that the symmetric error v_i dominates the one-sided error u_i. This implies that the disparity between the observed cost and the frontier cost is primarily due to factors outside the bank's control. On the other hand, large values of λ indicate that the one-sided error u_i dominates the symmetric error v_i and this means that the disparity between the observed cost and the frontier cost is mainly attributed to inefficiency. By testing the statistical significance of the λ parameter we can test whether any stochastic frontier specification, like (1), is required at all. In other words, if λ is not statistically different from zero, this would imply that σ_u^2 is zero and thus the one-sided error term u_i should be removed from the model and cost function (1) should be estimated using ordinary least squares (OLS).

Battese and Corra (1977), on the other hand, define a measure of the variation of cost from the frontier due to inefficiency, by analyzing the ratio of variability (γ) of u_i to total variability (u_i+v_i) in the following way:

$$\gamma = \frac{\sigma_u^2}{\sigma^2} \tag{5}$$

where $\sigma^2 = \sigma_v^2 + \sigma_u^2$. Consequently, $0 \leq \gamma \leq 1$. From (5), an estimate of γ is obtained using estimates of σ^2 and λ. The coefficient γ gives the percent of the discrepancy between the observed cost and the frontier one due to inefficiencies.

A criticism of the stochastic cost frontier approach is that the assumption of a half-normal distribution of inefficiencies is relatively inflexible. To address this criticism, the hypothesis of a half-normal distributed systematic error component is tested for the residuals of all banks over the period under examination. The test results show that the residuals are positively skewed, as implied by the assumption of an asymmetric half-normal distribution, for the systematic error component for all banks over the period under examination. In particular, the skewness measure for the entire period is 2.07 (p-value = 0.00).

3.2. *Model specification.* - It remains to choose a functional form for the cost frontier $\ln C(y_i, w_i)$. The translog functional form is used and is given below

$$\ln C = a_0 + \sum_i a_i \ln y_i + \sum_j b_j \ln w_j + \frac{1}{2}\sum_i \sum_j s_{ij} \ln y_i \ln y_j +$$

$$+ \frac{1}{2}\sum_i \sum_j g_{ij} \ln w_i \ln w_j + \sum_i \sum_j d_{ij} \ln y_i \ln w_j + k_1 t + k_2 t^2 + \sum_l f_l BN_l + \varepsilon \tag{6}$$

where: $s_{ij} = s_{ji}$ and $g_{ij} = g_{ji}$ by symmetry, and $\sum_j b_j = 1$, $\sum_j g_{ij} = 0 \ \forall i$, and $\sum_j d_{ij} = 0 \ \forall i$, by linear homogeneity, C is total cost, y_i is the quantity of output i, w_j is the price of input j, BN_l is a bank specific dummy and t is the time trend. We select the translog specification because of its flexibility. Moreover, it has been used extensively in many recent studies (Ferrier and Lovell, 1990; Berger, 1993; Allen and Rai, 1996; Lang and Welzel, 1996; Mester, 1996, 1997, among others).

In the banking literature there is some debate about what constitutes inputs and outputs for banks. Generally speaking, researchers appear to follow one of the two main approaches to the input and output specification, i.e. "the intermediation approach" and "the production approach". The intermediation

approach considers banks as financial intermediaries that convert deposits and purchased funds into loans and financial investments. This approach treats loans as outputs, while deposits and other liabilities are treated as inputs. Outputs are measured in value terms and costs include both interest expenses and production costs. By contrast, the production approach considers banks as producers of loan services and deposit accounts using capital and labor as inputs. Outputs are measured in terms of the number of accounts serviced and costs include production costs but do not include interest expenses. This study uses the intermediation approach, which is the approach most commonly used in the conventional cost function literature (Berger, 1993; Favero and Papi, 1995; Fukuyama, 1995; Allen and Rai, 1996; Lang and Welzel, 1996; Mester 1996, among others).

4. *Empirical Analysis*

4.1. *Data.* - Data for our analysis were obtained from the annual reports of 6 individual banks for the period 1982 to 1995. Four of the banks in the sample used are state banks, while the other two are private (see the Appendix). State banks control approximately 80% of the market and dominate the Greek banking industry. Private banks began really to play a substantial role in the banking sector over the post-1987 period.

This study specifies two types of output: traditional banking assets, such as loans (y_1), and investment assets (y_2). The output quantities are measured in value terms. Three inputs, whose prices are used to estimate the cost frontier, are considered: labor, capital and borrowed funds. The wage rate (w_1) is obtained by dividing staff expenses by total number of reported employees. The price of capital (w_2) is computed by dividing capital equipment and occupancy expenses (including depreciation, rental expenses, maintenance and repairs to property and equipment and property taxes) by fixed assets, net of depreciation. The borrowed funds price (w_3) is calculated by dividing total interest expenses (on demand and time deposits) by total interest liabilities. The cost variable comprises the sum of total interest expenses, staff expenses, and capital expenses.

4.2. *Cost frontier estimates.* - Maximum-likelihood (ML) estimates of the translog cost frontier (6) are given in Table 1. The estimated λ value is statistically different from zero, indicating that OLS is not the appropriate estimator in this study. The estimated value of λ is 2.1324, implying that the one-sided error u_i dominates the symmetric error v_i. Furthermore, the

Table 1. Maximum likelihood estimates of translog cost frontier (6)

Parameters-Variables			Estimates	Standard errors
a_0		cost intercept	42.4240	45.451
a_1		$\ln y_1$	0.8809	0.0788*
a_2		$\ln y_2$	2.8279	0.0876*
b_1		$\ln w_1$	–3.2102	4.2295
b_2		$\ln w_2$	–0.0237	0.0421
s_{11}	(1/2)	$\ln y_1 \ln y_1$	0.1585	0.0335*
s_{12}		$\ln y_1 \ln y_2$	–0.3728	0.1177*
s_{22}	(1/2)	$\ln y_2 \ln y_2$	0.3947	0.0889*
g_{11}	(1/2)	$\ln w_1 \ln w_1$	0.5199	0.5507
g_{12}		$\ln w_1 \ln w_2$	–0.2523	0.1204**
g_{22}	(1/2)	$\ln w_2 \ln w_2$	–0.0484	0.0675
d_{11}		$\ln y_1 \ln w_1$	0.2719	0.0357*
d_{12}		$\ln y_1 \ln w_2$	–0.0859	0.0072*
d_{21}		$\ln y_2 \ln w_1$	–0.2926	0.1017*
d_{22}		$\ln y_2 \ln w_2$	0.1704	0.0532*
k_1		t	–0.1131	0.0326*
k_2		t^2	0.0039	0.0018**
f_1		BN_1	0.9239	0.2639*
f_2		BN_2	0.9233	0.2367*
f_3		BN_3	0.3980	0.1730**
f_4		BN_4	1.9085	0.5022*
f_5		BN_5	0.4602	0.2091**
$\lambda = \sigma_u/\sigma_v$			2.1324	0.5340*
$\sigma=(\sigma_u^2+\sigma_v^2)^{1/2}$			0.4844	0.0806*
σ_u^2			0.1924	
σ_v^2			0.0423	
$\gamma = \sigma_u^2/\sigma^2$			0.8197	
No. of observations			84	
Adjusted R^2 on OLS			0.99264	
BP test			6.83 (p-value = 0.98)	
Log-likelihood			98.80	
Iterations completed			15	

Notes: y = banking output, measured as loans, w_1 = wage rate, w_2 = price of capital, t = time trend variable, *BNs* = bank specific dummies, *BP* is the Breusch-Pagan homoskedasticity test.
*, ** denote significance at 1% and 5% levels, respectively.

calculated γ value is 0.8197, suggesting that about 81.97% of the discrepancy between the observed cost and the frontier one is due to inefficiency. Thus, over the period 1982-1995, both λ and γ estimates imply that the divergence between the observed cost and the frontier one is primarily related to factors within the control of the banks, i.e. to X-inefficiency. In addition, Breusch-

Table 2. Frequency distribution of bank-specific X-inefficiencies based on the estimated cost frontier (6)

Year	1982	1983	1984	1985	1986	1987	1988
Inefficiency	Number of banks						
$0.0 \leq inef < 0.1$	0	0	0	0	0	0	0
$0.1 < inef < 0.2$	2	0	2	1	0	1	1
$0.2 \leq inef < 0.3$	2	3	2	2	4	1	2
$0.3 \leq inef < 0.4$	2	3	2	1	2	4	3
$0.4 \leq inef < 0.5$	0	0	0	2	0	0	0
$0.5 \leq inef < 0.6$	0	0	0	0	0	0	0
$0.6 \leq inef < 0.7$	0	0	0	0	0	0	0
$0.7 \leq inef < 0.8$	0	0	0	0	0	0	0
$0.8 \leq inef < 0.9$	0	0	0	0	0	0	0
$0.9 \leq inef < 1.0$	0	0	0	0	0	0	0
MEAN INEF.	0.2427	0.2944	0.3449	0.3105	0.2712	0.3025	0.2719
STD	0.0724	0.0593	0.1076	0.0923	0.0589	0.0716	0.0574
MIN	0.1526	0.2148	0.2086	0.1837	0.2050	0.1820	0.1941
MAX	0.3395	0.3662	0.4831	0.4193	0.3415	0.3920	0.3357
Year	**1989**	**1990**	**1991**	**1992**	**1993**	**1994**	**1995**
Inefficiency	Number of banks						
$0.0 \leq inef < 0.1$	0	0	0	0	1	0	0
$0.1 \leq inef < 0.2$	2	1	4	2	2	0	1
$0.2 \leq inef < 0.3$	2	4	1	2	0	3	3
$0.3 \leq inef < 0.4$	2	0	1	0	1	0	2
$0.4 \leq inef < 0.5$	0	1	0	2	0	2	0
$0.5 \leq inef < 0.6$	0	0	0	0	1	0	0
$0.6 \leq inef < 0.7$	0	0	0	0	1	1	0
$0.7 \leq inef < 0.8$	0	0	0	0	0	0	0
$0.8 \leq inef < 0.9$	0	0	0	0	0	0	0
$0.9 \leq inef < 1.0$	0	0	0	0	0	0	0
MEAN INEF.	0.2495	0.2619	0.1932	0.2781	0.3153	0.3908	0.2630
STD	0.0789	0.0893	0.0588	0.1171	0.2205	0.1456	0.0932
MIN	0.1355	0.1568	0.1390	0.1263	0.0814	0.2386	0.1231
MAX	0.3342	0.4184	0.3039	0.4255	0.6261	0.6158	0.3652

Pagan test for homoskedasticity indicates the absence of heteroskedasticity and the adequacy of the model.

For each single year of the period 1982-1995 the level of X-inefficiency of each individual bank is calculated by estimating the conditional expectation of the one-sided error component u_i from equation (4). Table 2 reports bank-

specific X-inefficiency measures, which are calculated for each year of the period under examination and are aggregated into a frequency distribution where the class interval is 0.1. In addition, the results show the mean level and range in the level of X-inefficiency of banks for each single year of the period 1982-1995. The 1991-year is the one with the lowest mean inefficiency level (i.e. 19.32%), while the 1994-year is the one with the highest mean inefficiency level (i.e. 39.08%). Furthermore, Table 2 shows that none of the banks in the sample is fully X-efficient. In particular, the year 1993 holds the banks with the lowest (i.e. 8.14%) and the highest (i.e. 62.61%) X-inefficiency levels of the entire time period under examination.

Table 3 presents some statistics of the calculated bank X-inefficiency levels for the entire period and for two sub-periods (i.e. 1982-1987 and 1988-1995). In particular, the mean bank inefficient level for the whole period is about 28.50%. This implies that, on average, 28.50% of the cost incurred over the period 1982-1995 can be avoided by the reorganization of the banking operation and without any loss in output. Furthermore, during the same period, the results show a wide variation of the inefficiency levels across banks; the minimum and maximum estimated inefficiencies are 8.14% and 62.61%, respectively. The period 1982-1995 is also divided into two sub-periods around the year 1987, because at that time, many regulatory changes started taking place in the Greek banking industry. Thus, it would be useful to compare bank inefficiencies between these two sub-periods. Table 3 shows that the mean bank X-inefficiency level of the 1982-1987 period (i.e. 29.44%) is greater than the one of the 1988-1995 period (i.e. 27.80%). However, the variation of the inefficiency levels across banks is lower in the first sub-period (i.e. between 15.26% and 48.31%) than in the second sub-period (i.e. between 8.14% and 62.61%). In other words, the frequency distributions of bank specific-inefficiencies of the 1988-1995 sub-period are more dispersed than the distributions of the 1982-1987 sub-period.

Table 3. X-inefficiency measures by period

Period	Mean	Std. Dev.	Minimum	Maximum
1982-1995	0.2850	0.1059	0.0814	0.6261
1982-1987	0.2944	0.0800	0.1526	0.4831
1988-1995	0.2780	0.1222	0.0814	0.6261

Overall, the Greek banking industry appears to display comparable average X-inefficiency levels relative to similar studies for other countries.

Based on the study by Allen and Rai (1996), average X-inefficiency estimates for small Spanish banks, small French banks, large Italian banks, large Japanese banks and large US banks are about 20.2%, 26.6%, 20.1%, 22.7% and 29.8%, respectively. However, the same study finds countries with much lower average levels of X-inefficiency than those given above (e.g. Australia 11.5%-13.4%, Austria 11.1%-10.3%, Canada 11.3%-12.8%, Switzerland 13%, Germany 12.3%-12.4%, Denmark 12.8%-14.5%, Sweden 13.3%-13.5%). Based on other stochastic frontier studies, average inefficiencies for Italian, Spanish, Tunisian, and US banks are about 30% (Resti, 1995), 15%-20% (Maudos, 1996a,b), 34%-39% (Chaffai, 1993), and 8%-15% (Mester, 1997), respectively.

Various bank specific factors could be responsible for the observed inefficiencies in the Greek banking sector. Such factors may be: total costs as a fraction of total assets (*COST*), total loans over total assets (*LOANS*), investment assets over total assets (*INVAS*), total deposits over total assets (*DEPOS*), bank size (*SIZE*) measured by total assets, interest expenses, staff expenses and capital expenses. The inefficiency measures can be examined by decomposing the bank-specific inefficiency levels against the above bank characteristics.

4.3. *Integration analysis.* - Unit root nonstationarity was tested by using the ADF test proposed by Dickey and Fuller (1981). The results are reported in Table 4. They show that the null hypothesis of stationarity was accepted only when first differences for total cost over total assets (*COST*), total loans over total assets (*LOANS*), investment assets over total assets (*INVAS*), total deposits over total assets (*DEPOS*), and total assets (*SIZE*) were used.

Table 4. Unit root test

		With trend	
Variable (X)		Levels	First differences
Cost/Assets	(*COST*)	–2.47(3)	–4.18(1)*
Loans/Assets	(*LOANS*)	–2.77(2)	–4.53(1)*
Inv. Assets/Assets	(*INVES*)	–3.35(3)	–8.08(3)*
Deposits/Assets	(*DEPOS*)	–2.80(2)	–4.69(2)*
Assets	(*SIZE*)	–2.49(2)	–5.03(1)*
Inefficiency	[$E(u_i\|\varepsilon)$]	–4.85(2)*	

Notes. Numbers in parentheses denote the number of lags in the augmented term of the test that ensures white noise residuals. These lags were determined using the Akaike criterion.
* denotes significance at 5%.

4.4. *Variance decompositions*. - This section determines quantitatively the degree of importance of various bank characteristics, which influence inefficiency beyond the sample period. The quantification takes place through variance decompositions. A Vector Autoregressive (VAR) model is developed which involves five (6) variables, i.e. the inefficiency measure $E(u_i|\varepsilon)$, total costs as a fraction of total assets (*COST*), total loans over total assets (*LOANS*), investment assets over total assets (*INVAS*), total deposits over total assets (*DEPOS*), and the size (*SIZE*) measured by total assets. Since in an unrestricted VAR model the ordering of the variables plays a significant role in the robustness of the results, by placing the inefficiency measure as a first variable, the ordering problem is extensively diminished. The presence of the variables in the VAR model is justified in Miller and Noulas (1997). In particular, the *LOANS* variable shows the liquidity risk for the banking sector, while the *SIZE* variable captures the presence of economies of scale in the sector. In addition, two dummy variables that determined the ownership (state = 1 or private = 0) of the bank and the bank specification bias were also employed. Since random innovations of the variables in the VAR model are capable of generating inefficiency, it is possible to decompose the total variance of inefficiency in each of the future periods and determine how much of this variance each variable explains. Therefore, the response of inefficiency to a one standard deviation innovation in each variable is obtained within a ten-year horizon. Since the innovations are not necessarily uncorrelated, the residual terms are orthogonalised using a Choleski decomposition in order to isolate the effects of each variable on inefficiency.

Table 5 reports the variance of the forecast error of inefficiency and its decomposition into proportions attributable to random innovation shocks to each variable, including its own. Panel I shows the decomposition results in the entire period under investigation, while panels II and III report the decomposition results over the regulated and the deregulated period, respectively.

The main results in Table 5 can be summarised as follows: Over the entire period under examination total cost as well as total loans (both measured as a fraction of total assets) explain about 60% of inefficiency (up to 2 years). Over a longer period (up to a 10-year horizon) the share of both of these variables drops to approximately 20%. Over the regulated period, it is the explanatory power of total loans that plays the major role in explaining inefficiency. In particular, within a 2-year forecasting horizon the explanatory power of loans amounts to 34%, while up to a 10-year horizon its explanatory power falls to 18%. Finally, over the deregulated period, it is the cost channel

that plays the major role in explaining banking efficiency. In particular, the contribution of this variable runs up to 47% over the first two years, while it falls to 32% within a 10-year forecasting horizon.

Table 5. Variance decompositions

Panel I. (Entire period: 1982-1995)						
		Variance error of inefficiency due to				
Forecasting horizon	$E(u_i\|\varepsilon)$	*COST*	*LOANS*	*DEPOS*	*SIZE*	*INVASS*
2	35.83	37.79	22.28	0.20	3.62	0.27
	(6.44)	(7.48)	(5.40)	(1.29)	(1.14)	(1.07)
4	37.29	26.33	23.36	4.06	4.64	4.32
	(8.31)	(7.04)	(8.71)	(1.53)	(1.35)	(1.74)
8	66.58	16.29	3.61	4.15	4.73	4.64
	(11.19)	(3.92)	(2.99)	(1.74)	(1.62)	(1.82)
10	66.57	16.28	3.61	4.16	4.74	4.64
	(10.76)	(3.84)	(2.73)	(1.68)	(1.57)	(1.66)

Panel II. (Regulated period: 1982-1987)						
		Variance error of inefficiency due to				
Forecasting horizon	$E(u_i\|\varepsilon)$	*COST*	*LOANS*	*DEPOS*	*SIZE*	*INVASS*
2	32.12	22.48	34.07	2.11	3.05	6.17
	(9.67)	(4.49)	(10.37)	(1.83)	(1.14)	(2.05)
4	28.84	22.92	26.88	14.10	2.73	4.53
	(9.45)	(4.86)	(8.91)	(2.11)	(1.37)	(2.00)
8	57.07	3.14	17.59	15.01	2.71	4.47
	(6.32)	(2.79)	(5.27)	(2.24)	(1.48)	(1.76)
10	57.01	3.19	17.59	15.03	2.71	4.47
	(7.08)	(3.04)	(5.32)	(2.19)	(1.39)	(1.63)

Panel III. (Deregulated period: 1988-1995)						
		Variance error of inefficiency due to				
Forecasting horizon	$E(u_i\|\varepsilon)$	*COST*	*LOANS*	*DEPOS*	*SIZE*	*INVASS*
2	25.86	47.00	20.30	2.54	2.47	1.84
	(7.76)	(8.04)	(5.81)	(1.16)	(1.46)	(1.09)
4	23.21	42.86	15.05	3.83	4.29	10.77
	(7.93)	(7.93)	(6.04)	(1.71)	(1.73)	(2.14)
8	49.92	31.61	5.33	4.96	6.51	1.66
	(6.81)	(6.84)	(5.79)	(1.82)	(1.92)	(2.27)
10	49.31	31.85	5.25	5.13	7.00	1.46
	(6.48)	(6.79)	(4.63)	(1.60)	(1.21)	(1.81)

Note: Numbers in parentheses denote standard errors estimated through Monte Carlo techniques with 500 replications.

5. *Conclusions and Policy Implications*

This paper uses the stochastic frontier approach to measure X-inefficiency of Greek banks. The empirical results show that the average X-inefficiency of Greek banks is about 28.50% for the 1982-1995 period. This indicates that, on average, 28.50% of the cost that incurred during this period can be avoided through the efficient use of existing inputs. Variance decomposition tests also revealed, at least over the period under examination, that the presence of operational inefficiency is mainly due to the mismanagement of loans and the critical presence of operating costs, though their contribution falls over a longer time period (up to 10 years).

These findings suggest that the personnel and administration costs have to be substantially reduced for the banking sector to guarantee its viability in the future. Moreover, in a rapidly increased competitive banking environment within the EU, certain actions have to be made that will lead to improved cost ratios and cost efficiency, such as mergers and acquisitions. Moreover, beginning in the period 1998-1999, commercial banks began to re-orient their asset structure towards more loans to the private sector and less investments in government bonds. At the same time, in conjunction with the rapid development of the capital market, they began to offer a larger variety of banking products and services, including brokerage services associated with stock trading in the domestic and international capital markets.

APPENDIX

State commercial banks: National Bank of Greece, Commercial Bank of Greece, Ionian Bank, and Bank of Macedonia and Thrace.

Private commercial banks: Alpha Credit Bank and Labour Bank.

REFERENCES

AIGNER D., LOVELL C.A.K., and SCHMIDT P., "Formulation and Estimation of Stochastic Frontier Production Function Models", *Journal of Econometrics*, 1977, *6*, pp. 21-37.

ALLEN L. and RAI A., "Economies of Scale and Scope in International Banking", *Journal of International Financial Markets, Financial Institutions and Money*, 1993. *3*, pp. 1-31.

—— and ——, "Operational Efficiency in Banking: An International Comparison", *Journal of Bankig and Finance*, 1996, *21*, pp. 1451-55.

BATTESE G.E. and CORRA G.S., "Estimation of a Production Frontier Model with Application to the Pastoral Zone of Eastern Australia", *Australian Journal of Agricultural Economics*,

1977, *21*, pp. 167-79.

BAUER P.W., BERGER A.N., and HUMPHREY D.B., "Efficiency and Productivity Growth in US Banking", in H.O. Fried, C.A.K. Lovell and S.S. Schmidt, eds., *The Measurement of Productive Efficiency: Techniques and Applications*, Oxford: Oxford University Press, 1993, pp. 386-413.

BERG S.A., FORSUND F.R., and JANSEN E.S., "Technical Efficiency of Norwegian Banks: A Nonparametric Approach to Efficiency Measurement", *Journal of Productivity Analysis*, 1991, *2*, pp. 127-42.

—— and KIM M., "Oligopolistic Interdependence and the Structure of Production in Banking: An Empirical Evaluation", *Journal of Money, Credit, and Banking*, 1994, *26*, pp. 309-22.

BERGER A.N., "Distribution-free Estimates of Efficiency in the U.S. Banking Industry and Tests of the Standard Distributional Assumptions", *The Journal of Productivity Analysis*, 1993, *4*, pp. 261-92.

—— and HUMPHREY D.B., "Efficiency of Financial Institutions: International Survey and Directions of Future Research", *European Journal of Operational Research*, 1997, *98*, pp. 175-212.

——, HUNTER W.C., and TIMME S.G., "The Efficiency of Financial Institutions: A Review and Preview of Research Past, Present, and Future", *Journal of Banking and Finance*, 1993, *17*, pp. 221-49.

CHAFFAI M.E., "Technical and Time Variant Allocative Inefficiency of Tunisian Commercial Banks: A Shadow Cost Frontier Approach Using Panel Data", Working Paper, Faculté des Sciences Economiques, Tunisia, 1993.

——, "Estimating Input-Specific Technical Inefficiency: The Case of the Tunisian Banking Industry", *European Journal of Operational Research*, 1997, *98*, pp. 315-32.

CHANG E.C., HASAN I., and HUNTER W.C., "Efficiency of Multinational Banks: An Empirical Investigation", Working Paper, New Jersey Institute of Technology, 1993.

DEBORGER B., FERRIER G.D., and KERSTENS K., "The Choice of a Technical Efficiency Measure on the Free Disposal Hull Reference Technology: A Comparison Using US Banking Data", Working Paper, University of Alaska, 1995.

DICKEY D.A. and FULLER W.A., "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root", *Econometrica*, 1981, *49*, pp. 1057-72.

DIETSCH M., "Les Surcapacités Bancaires en France", in D.E. Fair and R. Raymond, eds., *The Competitiveness of Financial Institutions and Centres in Europe*, Financial and Monetary Policy Studies, Dordrecht and Boston: Kluwer Academic on behalf of the Société Universitaire Européenne de Recherches Financières, vol 28, 1994, pp. 73-95.

EISENBEIS R.A., FERRIER G.D., and KWAN S.H., "An Empirical Analysis of the Informativeness of Programming and SFA Efficiency Scores: Efficiency and Bank Performance", Working Paper, University of North Carolina, Chapel Hill, NC, 1996.

FAVERO C.A. and PAPI L., "Technical Efficiency and Scale Efficiency in the Italian Banking Sector: A Non-parametric Approach", *Applied Economics*, 1995, 27, pp. 385-95.

FERRIER G.D. and LOVELL C.A.K., "Measuring Cost Efficiency in Banking: Econometric and Linear Programming Evidence", *Journal of Econometrics*, 1990, *46*, pp. 229-45.

FUKUYAMA H., "Measuring Efficiency and Productivity Growth in Japanese Banking: A Nonparametric Approach", *Applied Financial Economics*, 1995, *5*, pp. 95-117.

GREENWOOD J. and JOVANOVIC B., "Financial Development, Growth and the Distribution of Income", *Journal of Political Economy*, 1990, *98*, pp. 1076-1107.

JONDROW J., LOVELL C.A.K., MATEROV I.S., and SCHMIDT P., "On Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model", *Journal of Econometrics*, 1982, *19*, pp. 233-38.

KAPARAKIS E., MILLER S., and NOULAS A.G., "Short-Run Cost Inefficiency of Commercial Banks: A Flexible Stochastic Frontier Approach", *Journal of Money, Credit, and Banking*, 1994, *26*, pp. 875-93.

KWAN S.H. and EISENBEIS R.A., "An Analysis of Inefficiencies in Banking: A Stochastic Frontier Approach", Working Paper, Federal Reserve Bank of San Francisco, 1994.

LANG G. and WELZEL P., "Technology and Cost Efficiency in Universal Banking", Working Paper, University of Augsburg, Germany, 1995.

—— and ——, "Efficiency and Technical Progress in Banking: Empirical Results for a Panel of German Banks", *Journal of Banking and Finance*, 1996, *20*, pp. 1003-23.

LOZANO A., "Efficiency and Technical Change for Spanish Banks", Working Paper, University of Malaga. Spain, 1995.

MAUDOS J., (1996a), "A Comparison of Different Stochastic Frontier Techniques with Panel Data: An Application for Efficiency of Spanish Banks", Working Paper, University of Valencia, Spain. 1996.

——(1996b), "Market Structure and Performance in Spanish Banking Using a Direct Measure of Efficiency", Working Paper, University of Valencia, Spain, 1996.

MESTER L.J., "A Study of Bank Efficiency Taking into Account Risk-preferences", *Journal of Banking and Finance*, 1996, *20*, pp. 1025-45.

——, "Measuring Efficiency of US Banks: Accounting for Heterogeneity is Important", *European Journal of Operational Research*, 1997, *98*, pp. 230-43.

MILLER S.M. and NOULAS A.G., "The Technical Efficiency of Large Bank Production", *Journal of Banking and Finance*, 1996, *20*, pp. 495-509.

—— and ——, "Portfolio Mix and Large-Bank Profitability in the USA", *Applied Economics*, 1997, *29*, pp. 505-12.

PAGANO M., "Financial Markets and Growth: An Overview", *European Economic Review*, 1993, *37*, pp. 613-22.

RESTI A., "Linear Programming and Econometric Methods for Bank Efficiency Evaluation: An Empirical Comparison Based on a Panel of Italian Banks", Working Paper, University of Bergamo, Italy, 1995.

ROUBINI N. and SALA-I-MARTIN X., "Financial Repression and Economic Growth", *Journal of*

Development Economics, 1992, *39*, pp. 5-30.

SAINT PAUL G., "Technological Choice, Financial Markets and Economic Development", *European Economic Review*, 1992, *36*, pp. 763-81.

TAYLOR M.W., THOMPSON R.G., THRALL R.M., and DHARMAPALA P.S., "DEA\AR Efficiency and Profitability of Mexican Banks: A Total Income Model", *European Journal of Operational Research*, 1997, *98*, pp. 347-64.

TULKENS H., "On FDH Efficiency Analysis: Some Methodological Issues and Applications to Retail Banking, Counts, and Urban Transit", *Journal of Productivity Analysis*, 1993, *4*, pp. 183-210.

ZAIM O., "The Effect of Financial Liberalization on the Efficiency of Turkish Commercial Banks", *Applied Financial Economics*, 1995, *5*, pp. 257-64.

ABSTRACT

The goal of this study is to measure X-efficiency of the banking industry in Greece, through the methodology of a stochastic frontier. The empirical results show that Greek banks are characterized by a significant inefficiency measure, running up to virtually 29%, even following substantial deregulation activities in the monetary sector. In addition, error decomposition tests reveal that the presence of inefficiency is mainly attributed to high operating costs, i.e. administration costs, personnel costs, and loans monitoring costs.

JEL classification: *G21*
Keywords: *operational efficiency, banks, stochastic frontier, error decompositions, Greece*

RISEC, Volume 50 (2003), No. 4, 503-519

STOCK PRICE AND EXCHANGE RATE INTERACTIONS IN AN EMERGING MARKET

by
WAN N.W. AZMAN-SAINI*, MUZAFAR S. HABIBULLAH*, and MOHAMED AZALI*

1. *Introduction*

It is well documented that the recent Asian financial crisis spread from Thailand to the Philippines, South Korea, Malaysia and Indonesia. The impact on Malaysia showed that, between July 1997 and December 1997, the Malaysian ringgit (MYR) depreciated by 44.2% against the US dollar and reached the historic low of MYR 4.88 to a US dollar on January 7, 1998. From February through June 1998, the value of the ringgit was relatively stable between MYR 3.84 and MYR 3.98 per US dollar. From June to August 1998, the downward pressure on the ringgit intensified following the depreciation of the Japanese Yen, the contraction in the domestic economy and increased speculative activity. To stabilize the exchange rate, the initiative was taken on 1 September 1998 to impose selective exchange controls and to fix the exchange rate at MYR 3.80 to a US dollar on 2 September 1998. This represented 34% depreciation from its prevailing level prior to the crisis (Bank Negara Malaysia, 1999).

The immediate impact of the ringgit depreciation was on the Kuala Lumpur Stock Exchange (KLSE), with the Kuala Lumpur Composite Index (KLCI) declining by 44.8% in the second half of 1997.[1] During the period, it

* Department of Economics, Faculty of Economics and Management, Universiti Putra Malaysia, Serdang, Selangor (Malaysia). E-mail: muzafar@putra.upm.edu.my

The authors would like to thank an anonymous referee and the editor of this journal for helpful comments and suggestions on the earlier version of the paper. Any remaining errors are the solely responsibility of the authors.

[1] The KLCI is a broad-based capitalization-weighted index of 100 stocks designed to

is interesting to note distinct behaviors of the main and second board indices that may be indicative of the large and small firms stock prices respectively. As the Main Board (EMAS) index declined by 50% during the same period, the Second Board (SBIN) index declined by 71%.[2] Furthermore, as the reversal of short-term capital flows caused the KLCI to decline further and reached its lowest of 286 points on 1 September 1998, the EMAS index and SBIN index declined by another 45% and 50% respectively. Although it has been widely recognized that large firms have a higher degree of exchange rate exposures than small firms, this situation seems to indicate that small firms are more fragile than large firms.

The main purpose of this study is to empirically determine the causal relation between the exchange rate and stock prices in an emerging market, Malaysia. Specifically, our study focuses on the macro level issues by examining the causal behaviors of EMAS and SBIN indices. This study could contribute to the literature in the following ways. First, the study considers both bivariate and multivariate models. The majority of earlier empirical works have focused on the bivariate framework. It is shown that the lack of causal relationship in previous studies is due to the omission of important variables. It should be stressed that the use of an incomplete system that fails to account for other important variables may end up with spurious results. Second, we utilize the recent econometric methodology, which allows causal inferences to be conducted in systems including time series processes that may be integrated as well as cointegrated. The Granger non-causality procedure proposed by Toda and Yamamoto (1995) provides a simplistic approach in determining the causal relation between series regardless of the order of integration and/or the cointegration rank in a vector autoregressive (VAR) system. Finally, this study utilizes daily data to analyze the relationship between stock prices and exchange rate. It is well known that the use of low frequency observations may not be adequate to capture the dynamic relationship between exchange rates and stock prices. Hopefully, the use of daily data in this study can shed some light on the issue.

The motivation of our paper is to answer the following questions. What is the direction of causality between stock prices and the exchange rate? Has the relationship changed in recent years following the eruption of the Asian financial crisis in 1997? What has been the effect of this crisis on this

---

measure the overall performance of the Kuala Lumpur Stock Exchange.

[2] The EMAS and SBIN are the capitalization-weighted indices of all companies quoted and listed on the KLSE Main Board and Second Board respectively. For the listing purpose, firms must fulfill the paid-up capital requirement of MYR 60 millions for the main board and MYR 40 millions for the second board.

relationship? What is the impact of the US stock market on this relationship? Apart from providing insight into the theoretical matters of interest, such as informational efficient market hypothesis, the answers to these questions will have significant implications for understanding the relationship between exchange rate and stock prices in Malaysia.

The paper is organized as follows. Section 2 highlights some of the previous empirical findings. The data and estimating model are discussed in Sections 3 and 4 respectively. The results of the empirical test are discussed in Section 5. The final section concludes.

2. *Literature Review*

The Asian Financial Crisis and Economic Models. - Since the eruption of the financial crisis in Southeast Asia in the mid 1997, many economists have developed and used various economic models to explain the determinants of the crisis. These models basically can be classified as follows.

The "first-generation" model (FGM) explains that the crisis is caused by a loss of reserves to finance the budget deficit. The economy is under further pressure by the acceleration of the domestic credit expansion (see, Flood and Garber, 1984; Krugman, 1979). However, in the case of the 1997 crisis, the pre-crisis budget balances of the affected countries were either in surplus or had a limited deficit and thus the FGM seems unsuitable to be used in empirical works.

The "second-generation" model (SGM) exemplified by Obstfeld (1994) instead explains the crisis as a product of a conflict faced by the monetary authority to choose between maintaining a fixed exchange rate and an expansionary monetary policy. A crisis can be driven by a deterioration of domestic economic fundamentals or the result of self-validating shifts in expectation in the presence of multiple equilibrium conditions (see Cole and Kehoe, 1996; Sachs et al., 1996). The typical macroeconomic indicator considered in the literature focuses on real output growth, employment, and inflation (Corsetti et al., 1999b). However, the real GDP growth rates of the affected Asian economies prior to the 1997 crisis were very high and the rates of unemployment and inflation were low. Therefore, again this model does not seem fit to explain the crisis.

The failure of the FGM and SGM has led to the development of the "third-generation" model (TGM) that tries to explain the causes of the recent crisis. (see, Chang and Velasco, 1998; Corsetti et al., 1999a, 1999b; Goldstein, 1998; Krugman, 1998; Kwack, 1998; Letiche, 1998; Mishkin, 1996; Moreno

et al., 1998; Radelet and Sachs, 1998). Among these works, the most influential paper was that by Corsetti et al., (1999b) who consider moral hazard as the main cause of the recent Asian crisis. The model which is also known as "balance of payment crisis" model suggests that it was the government implicit guarantee that led to excessive external borrowing and eventual unsustainable current account deficits. The refusal of foreign creditors to refinance the cumulative loss left the local government with no choice other than to guarantee the external stock of liabilities. The expected inflationary fiscal reform (which possibly involves recourse to seigniorage revenue) eventually caused a collapse of the currency and led to the financial crisis.

Exchange Rates and Stock Prices. - Theoretically, there are two contrasting views on the linkages between stock prices and exchange rates. On one hand, the proponents of the 'portfolio-balance' model believe that stock prices negatively affect exchange rates. Equities being part of wealth, may affect exchange rates through the demand for money. For example, higher stock prices may result in a higher demand for money with ensuing higher interest rates. This, in turn, would attract foreign capital flows to the country and result in the appreciation of the domestic currency.

On the other hand, another group of economists believe that the currency appreciation under the floating exchange rate regime would affect the international competitiveness of local products and the trade balance position of the nation.[3] Eventually the firms' future cash flows are affected due to a real output deterioration and this will result in lower stock prices. Generally, this model suggests that exchange rates positively affect stock prices.

Earlier empirical researches (see, Ang and Ghalap, 1976; Frank and Young, 1972) into gauging the linkages between exchange rates and stock prices in developed markets (especially the US market) have mainly focused on the exposure of Multinational Corporations (MNCs) stock prices to the exchange rate volatility. For example, Frank and Young (1972) examination of the US MNCs found that there is no recognizable pattern of stock price reactions to the exchange rate. Complementing this study, Ang and Ghalap (1976) who examined fifteen US MNCs conclude that the stock market is efficient and adjusts rapidly to the changes of the exchange rate. However, the empirical examinations of the relation between stock returns and exchange rates during the flexible exchange regime have largely provided a mixed evidence (see, Bahmani-Oskooee and Sohrabian, 1992; Mok, 1993;

[3] For example. the 'flow-oriented' model by Dornbusch and Fisher (1980).

Smith, 1992). For example, Bahmani-Oskooee and Sohrabian (1992) conclude that there is a feedback interaction between US stock prices and exchange rates. This finding is consistent with Mok's (1993) study on Hong Kong. In contrast, Smith (1992) found that stock returns have significant influence on exchange rates in Germany, Japan and the United States. In addition, Granger et al. (2000) found that exchange rates lead stock prices in Japan and Hong Kong for the period January 95 to November 97 and January 86 to November 87 respectively.

In the case of emerging markets, Abdalla and Murinde (1997) investigate the interactions between exchange rates and stock prices in the financial markets of India, South Korea, Pakistan and the Philippines. Employing the monthly data in a form of bivariate vector autoregressive (VAR) model, the analysis over the period from January 1985 to July 1994 supports the "flow-oriented" model according to which exchange rate changes affect stock prices in India, Pakistan and South Korea. In the case of the Philippines market, stock prices take the lead. In contrast, Granger et al. (2000) using daily data over the period January 1987 to December 1994 found that the exchange rate leads stock prices and a bi-directional causality exists in the Korean financial market for January 1995 to November 1997 period. Furthermore, the feedback interaction also exists for data from Indonesia, whereas the result of the Taiwan market indicates that stock prices contain leading information.

With respect to Malaysia, several studies have documented mixed results. For example Ajayi et al. (1998) found that there is no causal relationship between the exchange rate and stock prices between December 1987 and September 1991. Consistent with this result, Granger et al. (2000) fail to unveil any significant lead-lag relationship for the sample period that span from January1986 to November 1987. In contrast, Ibrahim (2000) uncovers a bi-directional causality during the sample period from January 1979 to June 1996. This finding is consistent with Granger et al. (2000) for the different sub-sample period from January 1995 to November 1997. However, the results for the period from December 1987 to December 1994 support the existence of unidirectional causality from the exchange rate to stock prices.

In other related developments, using simple multiple regression analysis, Amain and Law (2000) conclude that the exchange rate contains a predictive power about future stock return movements. This evidence is particularly significant during the weak ringgit period (i.e. July 1997 until July 1998). Ten sectoral equity indices were utilized over the period from September 1993 to July 1998 employing weekly closing prices. As for the strong ringgit period (i.e. September, 1993 to January, 1997), there is no definite relation between the exchange rate and stock prices. However, the application of

regression analysis invalidates the possibility of testing a causal relationship.

As far as the model specification is concerned, the majority of the previous macro-level empirical works have focused on a simple bivariate framework. It should be stressed that the application of a simple bivariate model in the causality test is subject to possible mispecification bias. The omission of important variables may lead to type 'I' and type 'II' errors, that is, spurious rejection of causal relationship as well as spurious detection of it. In a macroeconomic framework, the relationship between exchange rate and stock prices can best be captured by including other variables (see Smith, 1992).[4] Therefore, in addition to the bivariate analysis, we also analyze the relationship under the multivariate setting by allowing US stock prices to exert their influence. We trust that the inclusion of US stock prices in the system, which have the greatest impact on world markets, will shed some light on the issue.

3. *Data*

Daily data set for exchange rate and stock prices covering January 1991 to August 1998 were utilized.[5] The advantage of using daily data is that they may better capture the dynamics of the interaction between equities and exchange rates, whereas any significant interaction between the two series may be diluted in data of low frequency. The starting date is dictated by the availability of the exchange rate data. The period after August 1998 is avoided due to the imposition of capital controls which completely closed the financial market to foreigners. End of day spot bilateral exchange rates vis-à-vis the US dollar are utilized.[6] The capitalization-weighted Main Board (EMAS) and Second Board (SBIN) indices are used to proxy the Malaysian stock prices. The S&P 500 index represents the US stock market. In order to better understand the relationship between stock prices and the exchange rate, the sample period is divided into two sub-samples. The pre-crisis sub-sample period covers from January 1991 to June 1997 and the crisis period continued from July 1997 through August 1998. All data were collected from various issues of the *Investor Digest* and *Monthly Statistical Bulletin* published by Kuala Lumpur Stock Exchange (KLSE) and Bank Negara Malaysia (BNM) respectively. Prior to the analysis, the data were converted into natural

[4] Smith (1992) includes data on current account surplus, interest rate differential, bond prices and debt.

[5] The starting date for EMAS index is November 1, 1991.

[6] The exchange rate is expressed in Malaysian ringgit per unit of US dollar.

logarithm.

4. *The Estimating Model*

It is well known that in the context of integrated series, the conventional application of the *F*-test (i.e. in a standard VAR model) is unsuitable.[7] Moreover, the *F*-test is not valid unless the variables are cointegrated in levels. Recently, an alternative approach which utilizes the Modified *WALD* (*MWALD*) test for testing linear restriction on the parameters was proposed by Toda and Yamamoto (1995). This test has an asymptotic χ^2 distribution when a VAR ($k + d_{max}$) is estimated where d_{max} is the maximum degree of integration suspected to occur in the system and k is the optimal lag length. Therefore, the Toda-Yamamoto non-causality procedure is viewed as a long-run causality test. Toda and Yamamoto (1995) point out that, for $d = 1$, the lag selection procedure is always valid since $k \geq 1 = d$. If $d = 2$, the procedure is valid unless $k = 1$. Moreover, the *MWALD* statistic is valid regardless whether a series is $I(0)$, $I(1)$ or $I(2)$, non-cointegrated or cointegrated of an arbitrary order. In addition, Zapata and Rambaldi (1997) provide evidence that the *MWALD* test has a comparable performance in size and power to the LR and *WALD* tests if (i) the correct number of lags for estimating $k + d_{max}$ is identified and (ii) no important variables are omitted, provided a sample of 50 observations is available.

Rambaldi and Doran (1996) have demonstrated that the *MWALD* procedure for testing Granger non-causality can be easily constructed using a Seemingly Unrelated Regression (SUR). Following Toda and Yamamoto (1995) Granger non-causality test, these variables can be causally linked in a system of two variables, with 3-order VAR as follows[8]:

$$(1) \quad \begin{bmatrix} SP_t \\ EX_t \end{bmatrix} = \alpha_0 + \alpha_1 \begin{bmatrix} SP_{t-1} \\ EX_{t-1} \end{bmatrix} + \alpha_2 \begin{bmatrix} SP_{t-2} \\ EX_{t-2} \end{bmatrix} + \alpha_3 \begin{bmatrix} SP_{t-3} \\ EX_{t-3} \end{bmatrix} + \alpha_4 \begin{bmatrix} \varepsilon_{SP} \\ \varepsilon_{EX} \end{bmatrix}$$

where α_0 is an identity matrix and $E(\varepsilon_t) = [\varepsilon_{SP_t} \varepsilon_{EX_t}]' = 0$ and $E(\varepsilon_t \varepsilon_t') = \Sigma$. For example, if $k = 2$ and $d_{max} = 1$, a causality from *EX* to *SP* can be established by rejecting the null hypothesis that EX_{t-1} and EX_{t-2} are jointly equal to zero in the first equation of the above system where EX_{t-3} is left

[7] For spurious regression due to non-standard distribution of the test statistics, see Toda and Phillips (1993).

[8] *SP* = stock prices; *EX*= exchange rates.

unrestricted as a long run correction mechanism. A similar procedure can be used to test the causality from *SP* to *EX* by establishing a significance of the *MWALD* statistic for testing that a group of lagged *SP* variables are jointly equal to zero in the second equation of the system. Further discussions on this procedure appear in the Appendix.

5. *The Empirical Results*

Prior to testing the non-causality test, it is necessary to determine the degree of integration of stock market indices and exchange rates. The two commonly used tests of unit root are Dickey-Fuller (1979) (DF) and Phillips-Perron (1988) (PP) tests. In both tests, the null hypothesis is that a unit root exists in the autoregressive representation of the time series. The DF type test attempts to account for temporally dependent and heterogeneously distributed errors by including lagged sequences of first differences of the variable in its set of regressors. The Augmented Dickey-Fuller (ADF) test (see, Said and Dickey, 1984) extends the DF test by allowing a higher order of autoregressive process. The semi-parametric PP type tests try to account for dependent and IID processes through adopting a non-parametric adjustment, hence eliminating any nuisance parameters.

However, as evident in DeJong et al. (1992) and Schwert (1987), these tests tend to accept the null of unit root too frequently against a stationary alternative due to lack of power. Therefore, the failure to reject a unit root may be simply due to the low power of the standard unit root test (Masih and Masih, 1999). In addition, Stock (1995) points out that nuisance parameters such as the largest autoregressive root are quite typical of financial time series. Due to the growing controversy surrounding the specific test to employ, we complement the ADF and PP tests with the mean stationary test proposed by Kwiatkowski et al. (1992) (KPSS). The KPSS test is based on the following statistic

$$(2) \qquad \eta(u) = (1/T^2)\sum_{t=1}^{T} S_t^2 / \sigma_k^2, \quad \text{where} \quad S_t = \sum_{i=1}^{T} v_t, \; t = 1,.....T$$

where v_t is the residual term from a regression of y_t on an intercept, and σ_t^2 is a consistent long-run estimate of y_t, and T represents the sample size. If the computed value of $\eta(u)$ is larger than the critical value, then the null hypothesis of stationarity is rejected. We consider a variable to contain a unit

root or be unit-root non-stationary if the null hypothesis of unit root is not rejected by the ADF and PP tests but the null hypothesis that the variable is mean stationary is rejected by the KPSS test.

The results of unit root tests reported in Table 1 indicate that the level of each series is nonstationary. Using the ADF, PP and KPSS tests with trend, we found that for all series the hypothesis of level non-stationarity cannot be rejected. However, results for first-difference clearly show that the hypothesis of non-stationarity can be rejected. Thus, like most financial time series, the stock prices and exchange rate require differencing to achieve stationarity or they are $I(1)$.

Table 1. Results of unit root tests

Period	Variables	ADF*(lag)	PP*(lag)	KPSS#(lag)	Concluding remarks
			Level		
Pre–crisis	*EX*	−1.8524 (4)	−1.8718 (7)	**0.2551** (21)	Non-stationary
	EMAS	−1.3453 (4)	−1.2916 (7)	**0.7734** (21)	Non-stationary
	SBIN	−2.1918 (4)	−2.0569 (7)	**0.2854** (21)	Non-stationary
	SP500	−1.0510 (4)	−0.9822 (7)	**0.9847** (21)	Non-stationary
Crisis	*EX*	−2.2190 (4)	−2.2447(5)	**0.2752** (20)	Non-stationary
	EMAS	−1.8515 (4)	−1.7970(5)	**0.1443** (20)	Non-stationary
	SBIN	−2.5367 (4)	−2.1928(5)	**0.1242** (20)	Non-stationary
	SP500	−1.0170 (4)	−1.5379(5)	**0.1421** (20)	Non-stationary
			First Difference		
Pre–crisis	*EX*	**−19.0102** (4)	**−40.8942** (7)	0.0432 (21)	Stationary
	EMAS	**−15.8225** (4)	**−38.4827** (7)	0.0617 (21)	Stationary
	SBIN	**−15.0226** (4)	**−34.4558** (7)	0.0491 (21)	Stationary
	SP500	**−18.4149** (4)	**−39.3599** (7)	0.1163 (21)	Stationary
Crisis	*EX*	**−8.6458** (4)	**−16.6262** (5)	0.0613 (20)	Stationary
	EMAS	**−7.9801** (4)	**−15.5227** (5)	0.1090 (20)	Stationary
	SBIN	**−7.7542** (4)	**−13.2473** (5)	0.0504 (20)	Stationary
	SP500	**−8.3893** (4)	**−18.2799** (5)	0.1134 (20)	Stationary

Notes: Boldface numbers indicate rejection of null hypothesis at the 10% level of significance.
* The test statistics were computed with lag not exceeding the floor function $floor(4(T/100)^{2/9})$.
# Lag lengths were determined based on $[cT]^k$, where $c = 5$ and $k = 0.25$ are adopted following Bierens (2001).

Having identified that all series exhibit $I(1)$ behavior or the maximum degree of integration in the system is one, we examine the nature of causal linkages between the series by performing the Granger non-causality test. Applying the methodology proposed by Toda and Yamamoto (1995) outlined

in Section 4, we examine the nature of causal linkages between stock prices and the exchange rate. The bivariate model was estimated and the results for both sub-periods are reported in Table 2.

Table 2. Results of bivariate Granger non-causality tests due to Toda and Yamamoto (1995)

	Hypothesis:	MWALD	(lag)	Concluding remarks
Panel A: Pre-Crisis	*EMAS* $-/\rightarrow$ *EX*	**68.6698**	(13)	*EMAS* $\longrightarrow$ *EX*
	EX $-/\rightarrow$ *EMAS*	**112.1646**	(13)	*EX* $\longrightarrow$ *EMAS*
	SBIN $-/\rightarrow$ *EX*	**75.1966**	(10)	*SBIN* $\longrightarrow$ *EX*
	EX $-/\rightarrow$ *SBIN*	**79.8818**	(10)	*EX* $\longrightarrow$ *SBIN*
Panel B: Crisis	*EMAS* $-/\rightarrow$ *EX*	**200.3299**	(2)	*EMAS* $\longrightarrow$ *EX*
	EX $-/\rightarrow$ *EMAS*	**56.2334**	(2)	*EX* $\longrightarrow$ *EMAS*
	SBIN $-/\rightarrow$ *EX*	**321.9195**	(8)	*SBIN* $\longrightarrow$ *EX*
	EX $-/\rightarrow$ *SBIN*	**157.0801**	(8)	*EX* $\longrightarrow$ *SBIN*

Notes: Boldface numbers indicate rejection of null hypothesis at the 10 % level of significance.
($-/\rightarrow$) and ($\longrightarrow$) indicate no Granger-cause and Granger-cause respectively.
Lag lengths were determined using Akaike Information Criterion (AIC).

Panels A and B show the results for pre-crisis and crisis periods respectively. The results of the bivariate model for both EMAS and SBIN indices suggest that there exists a bi-directional causality between exchange rate and stock prices (i.e. the null of Granger non-causality is rejected at the 10% level). However, as explained earlier, the bivariate model may be flawed due to the omission of important variables. Therefore, we re-examine the causal relationship between stock prices and exchange rate by including US stock prices (i.e. S&P 500 index) in Equation 1. The US stock prices are selected due to the important impact the US market has had on Asian equity market.[9] In the present context, the US stock market, which represents the world market, could be a channel through which the foreign exchange and stock are linked. The results of this estimation for both sub-samples are reported in Table 3.

[9] Wei et al. (1995) found that the US equity market has more impact than Japan's on the Asian equity markets.

Table 3. Results of multivariate Granger non-causality tests due to Toda and Yamamoto (1995)

	Hypothesis:	MWALD	(lag)	Concluding remarks
Panel A: Pre-Crisis	$EMAS \not\rightarrow EX$	**9.6707**	(5)	$EMAS \rightarrow EX$
	$EX \not\rightarrow EMAS$	**31.4393**	(5)	$EX \rightarrow EMAS$
	$SBIN \not\rightarrow EX$	0.4746	(5)	$SBIN \not\rightarrow EX$
	$EX \not\rightarrow SBIN$	**41.7025**	(5)	$EX \rightarrow SBIN$
Panel B: Crisis	$EMAS \not\rightarrow EX$	**12.3234**	(2)	$EMAS \rightarrow EX$
	$EX \not\rightarrow EMAS$	2.3298	(2)	$EX \not\rightarrow EMAS$
	$SBIN \not\rightarrow EX$	**15.3622**	(2)	$SBIN \rightarrow EX$
	$EX \not\rightarrow SBIN$	**22.8860**	(2)	$EX \rightarrow SBIN$

Notes: Boldface numbers indicate rejection of null hypothesis at the 10 % level of significance.
($\not\rightarrow$) and ($\rightarrow$) indicate no Granger-cause and Granger-cause respectively.
Lag lengths were determined using Akaike Information Criterion (AIC).

With respect to the EMAS index, the main board index of the KLSE, the results for the pre-crisis period suggest that both the null hypothesis of Granger non-causality from stock prices to exchange rate and from exchange rate to stock prices can be rejected at the 10% level. The inclusion of US stock prices apparently does not affect the lead-lag structure for the pre-crisis period. This indicates that there is strong evidence to support a bi-directional causality or feedback interaction before the crisis unfolded. However, the result for the crisis period indicates that only the null hypothesis of Granger non-causality from exchange rate to stock prices can be rejected at the 10% significant level. This is indicative of the one-way causality running from exchange rate to stock prices of the main board.

On the other hand, the examination of the SBIN index for the pre-crisis period suggests that inclusion of US stock prices does affect the lead-lag structure as there is only one-way causality running from stock prices to exchange rate. Only the null of Granger non-causality running from stock prices to exchange rate was rejected at the 10 percent significance level. However, the results for the crisis sub-sample period uncover that both nulls were rejected at the 10 percent significant level. This is consistent with the finding tested on the bivariate framework. Thus, there is a strong evidence to support the existence of a feedback interaction between SBIN index and exchange rate during the recent crisis.

6. *Conclusion*

This study contributes to the literature on the causal relationship between stock prices and exchange rates by examining Malaysian daily data as a case study. Specifically, this study analyzes both the KLSE main and second board indices. We employ both bivariate and multivariate models covering a sample period from January 1991 to August 1998. Essentially, the unit root properties of each series were determined using the ADF, PP and KPSS tests. Having established the order of integration of the individual series, we move to testing for causal linkages using the Granger non-causality test prescribed by Toda and Yamamoto (1995). Among the findings of interest, the bivariate analysis suggests bi-directional causal relations for all cases. However, this model is potentially misspecified and may be flawed due to omission of variables. Therefore, we re-examine the relationship between stock prices and exchange rates in a multivariate setting by allowing US stock prices to exert their potential influence on local stock prices and exchange rates.

The result for the pre-crisis period indicates that the inclusion of US stock prices does not change the lead-lag structure between EMAS index and exchange rate. The bi-directional causality prevails between the series. With respect to the informational efficient market hypothesis, this finding suggests that the prices of the stocks listed on the main board and foreign exchange markets are not efficient with respect to each other as information in one market can be used to predict the movement of other prices. However, the result for the SBIN index suggests that there is only one-way causality running from exchange rate to stock prices. This is consistent with the flow-oriented model which means that the exchange rate movements could be used to study the stock price movements.

With respect to the crisis period, the result is indicative of unidirectional causality from EMAS index to exchange rate. This is consistent with the portfolio-balance model which stresses the important impact of free capital flows. It has been recognized that the free flows of capital can be beneficial as well as hazardous to the financial system. However, the inclusion of the US stock prices does not affect the lead-lag structure between SBIN index and exchange rate. Thus, there is strong evidence to support the bi-directional causality between the series. In this case, both the flow-oriented and portfolio balance models run their courses.

In conclusion, future study may include other control variables such as the Japanese Stock Index as well as the influence of the yen on the Malaysian ringgit. We leave this for future research.

APPENDIX

Granger non-causality test due to Toda and Yamamoto (1995)

Toda and Yamamoto (1995) proved that in integrated and cointegrated systems the Wald statistic for testing linear restrictions on the parameters of VAR(k) has an asymptotic chi-square (χ^2) distribution when a VAR ($k + d_{max}$) is estimated, where d_{max} is the maximum order of integration in the system and k is optimal lag length. Basically, Toda and Yamamoto proposed an Ordinary Least Square (OLS) estimation of VAR in level of the equation (A.1).

$$(A.1) \quad y_t = \lambda_0 + \lambda_1 t + ... + \lambda_q t^q + \theta_1 y_{t-1} + ... + \theta_k y_{t-k} + ... + \theta_p y_{t-p} + \varepsilon_t$$

where t is time trend (i.e. $t = 1, ..., T$) and $p \geq (k + d_{max})$ consisting of y_t's that are $I(d)$ which may be cointegrated or not, θ's are matrices of coefficient and ε_t is the error term with mean zero and variance-covariance matrix Σ. The hypothesis of restrictions under test will assume the terms $\theta_{k+1}, ... \theta_p$ to be zero and thus ignored. In the usual matrix notation, equation (A.1) may be expressed as:

$$(A.2) \quad Y' = \Xi\Lambda' + \Theta X' + \psi Z' + E'$$

where $\Xi = [\lambda_0, ...\lambda_q]$, $\Lambda = [\tau_1, ..., \tau_T]$ with $\tau_t = (1, t, ..., t^q)'$, $\Theta = [\theta_1, ..., \theta_k]$, $X_t = [x_1, ..., x_T]$ with $x_t = (y'_{t-1}, ..., y'_{t-k})'$, $\psi = [\theta_{k+1}, ..., \theta_p]$ and $Z = [z_1, ..., z_T]$ with $z_t = (y'_{t-k-1}, ..., y'_{t-p})'$, and $E = (\varepsilon_1, ..., \varepsilon_T)$.

The null hypothesis of Granger non-causality (i.e. H_0: $f(\phi) = 0$ and where ϕ is vector (Θ)) is tested by the following Wald statistic (w):

$$(A.3) \quad w = \frac{f(\hat{\phi})'}{F(\hat{\phi})\{\Sigma_\varepsilon \otimes (X'\hat{Q}X)^{-1}\}F(\hat{\phi})'} f(\hat{\phi})$$

where $\Sigma_\varepsilon = T^{-1}\hat{E}'\hat{E}$, $\hat{Q} = \hat{Q}_\tau - \hat{Q}_\tau Z(Z'\hat{Q}_\tau Z)^{-1} Z'\hat{Q}_\tau$, and $\hat{Q}_\tau = I_T - \hat{\Lambda}(\hat{\Lambda}'\hat{\Lambda})^{-1}\hat{\Lambda}'$. If the computed statistic (w) is greater than the critical value, the null is rejected implying Granger causality and vice versa.

The main advantage of this methodology lies in its theoretical simplicity and ability to overcome many of the shortcomings associated with other procedures especially the cointegration test of Johansen (1988) and Johansen and Juselius (1990). There is growing concern that the LR tests of Johansen (1988) and Johansen and Juselius (1990) have often not provided the degree of empirical support that might reasonably have been expected for a long run relationship (Shan et al., 1997). In addition, simulation studies by Bewly and Yang (1996) and Ho and Sorensen (1996) show that the power of LR tests deteriorates rapidly with the augmentation of optimal lag length. Another related advantage of this method is that it avoids the

difficulty of estimating causal relationship when more than one cointegrating vectors exist.

REFERENCES

ABDALLA I.S. and MURINDE V., "Exchange Rate and Stock Price Interactions in Emerging Financial Markets: Evidence on India, Korea, Pakistan, and Philippines", *Applied Financial Economics*, 1997, *7*, pp. 25-35.

AJAYI R.A., FRIEDMAN J. and MEHDIAN S.M., "On the Relationship between Stock Returns and Exchange Rates: Tests of Granger Causality", *Global Finance Journal*, 1998, *9*(2), pp. 241-51.

AMAIN B.A. and LAW S.H., "The Relationship between Stock Price and Exchange Rate: Empirical Evidence Based on the KLSE Market", *The Asian Economic Review*, 2000, *42*, pp. 39-49.

ANG J.S and GHALAP A., "The Impact of US Devaluation on the Stock Prices of Multinational Corporations", *Journal of Business Research*, 1976, *4*, pp. 25-34.

BAHMANI-OSKOOEE M. and SOHRABIAN A., "Stock Prices and the Effective Exchange Rate of the Dollar", *Applied Economics*, 1992, *24*, pp. 459-64.

BANK NEGARA MALAYSIA, *The Central Bank and Financial System in Malaysia – A Decade of Change*, Kuala Lumpur, 1999.

——, *Monthly Statistical Bulletin*, various issues.

BEWLY R. and YANG M., "On the Size and Power of System Tests for Cointegration", in M. McAleer et al., eds., *Proceedings of the Econometric Society Australasian Meeting*, Australia: Perth, 1996, pp. 1-20.

BIERENS H.J., "EasyReg International 2001", Department of Economics, Pennsylvania State University, University Park, PA, 2001.

CHANG R. and VELASCO A., "The Asian Liquidity Crisis", Working Paper Series No. 98–11, Federal Reserve Bank of Atlanta, July 1998.

COLE H. and KEHOE T., "A Self-Fulfilling Model of Mexico's 1994-1995 Debt Crisis", *Journal of International Economics*, 1996, *41*, pp. 309-30.

CORSETTI G., PESENTI P., and ROUBINI N. (1999a), "What Caused the Asian Currency and Financial Crisis?", *Japan and the World Economy*, 1999, *11*(3), pp. 305-73.

—— (1999b), "Paper Tigers? A Model of the Asian Crisis", *European Economic Review*, 1999, *43*, pp. 1211-36.

DEJONG D.N., NANKERVIS J.C., SAVIN N.E., and WHITEMAN C.H., "Integration versus Trend Stationarity in Time Series", *Econometrica*, 1992, *60*, pp. 423-33.

DICKEY D.A. and FULLER W.A., "Distribution of the Estimators for Autoregressive Time Series with a Unit Root", *Journal of American Statistical Association*, 1979, *74*, pp. 427-31.

DORNBUSCH R. and FISHER S., "Exchange Rates and the Current Account", *American Economic Review*, 1980, *70*, pp. 960-71.

FLOOD R. and GARBER P., "Collapsing Exchange–Rate Regimes: Some Linear Examples", *Journal of International Economics*, 1984, *17*, pp.1-13.

FRANK P. and YOUNG A., "Stock Price Reaction of Multinational Firms to Exchange Realignments", *Financial Management*, 1972, *1*, pp. 66-73.

GOLDSTEIN M., "The Asian Financial Crisis: Causes, Cures, and Systemic Implications", Policy Analyses in International Economics 55, Institute for International Economics, 1998.

GRANGER C.W.J.. BWO-NUNG H., and YANG C.W., "A Bivariate Causality between Stock Prices and Exchange Rates: Evidence from Recent Asia Flu", *The Quarterly Review of Economics and Finance*, 2000, *40*, pp. 337-54.

HO M.S. and SORENSEN B.E., "Finding Cointegration Rank in High Dimensional Systems Using the Johansen Test – An Illustration Using Data Based on Monte Carlo Simulation", *Review of Economics and Statistics*, 1996, *78*(4), pp. 726-32.

IBRAHIM M., "Cointegration and Granger Causality Test of Stock Prices and Exchange Rate Interaction in Malaysia", *ASEAN Economic Bulletin*, 2000, *17*(1), pp. 36-46.

JOHANSEN S., "Statistical Analysis of Cointegrating Vectors", *Journal of Economics Dynamics and Control*, 1988, *12*, pp. 231-54.

—— and JUSELIUS K.. "Maximum Likelihood Estimation and Inference on Cointegration, with Applications to the Demand for Money", *Oxford Bulletin of Economics and Statistics*, 1990, *52*, pp.169-200.

KRUGMAN P., "A Model of Balance of Payment Crises", *Journal of Money, Credit and Banking*, 1979, *11*, pp. 311-25.

——, "What Happened to Asia?", mimeo, Massachusetts Institute of Technology, 1998.

KUALA LUMPUR STOCK EXCHANGE, *Investor Digest*, various issues.

KWACK S.Y.. "Factors Contributing to the Financial Crisis in Korea", *Journal of Asian Economics*, 1998, *9*(4), pp. 611-25.

KWIATKOWSKI D.. PHILLIPS P.C.B., SCHMIDT P., and SHIN Y., "Testing the Null Hypothesis Stationarity against the Alternative of a Unit Root: How Sure Are We that Economic Time Series Have a Unit Root?", *Journal of Econometrics*, 1992, *54*, pp. 159-78.

LETICHE J.M., "Causes of the Financial and Economic Crisis in Southeast Asia and the Need for National, Regional, and IMF Structural Reform", *Journal of Asian Economics*, 1998, *9*(2), pp. 181-91.

MASIH A.M.M. and MASIH R., "Are Asian Stock Market Fluctuations Due Mainly to Intra-Regional Contagion Effects? Evidence Based on Asian Emerging Stock Markets", *Pacific-Basin Finance Journal*, 1999, *7*, pp. 251-82.

MISHKIN F.S., "Understanding Financial Crises: A Developing Country Perspective", *NBER Working Paper* No. 5600, May 1996.

MOK H.M.K., "Causality of Interest Rate, Exchange Rate and Stock Price at Stock Market

Open and Close in Hong Kong", *Asia Pacific Journal of Management*, 1993, *10*, pp. 123-43.

MORENO R., PASADILLA G., and REMOLONA E., "Asia's Financial Crisis: Lessons and Policy Responses", Pacific Basin Working Paper No. PB98–02, Federal Reserve Bank of San Francisco, July 1998.

OBSTFELD M., "The Logic of Economic Crises", *Cahiers Economiques et Monétaires*, 1994, *43*, pp. 189-213.

PHILLIPS P.C.B. and PERRON P., "Testing for Unit Root in Time Series Regression", *Biometrika*, 1988, *75*, pp. 335-46.

RADELET S. and SACHS J., "The East Asian Financial Crisis: Diagnosis, Remedies, Prospects", *Brookings Papers on Economic Activity*, 1998, *1*, pp. 1-74.

RAMBALDI A.N and DORAN H.E., "Testing for Granger Non-Causality in Cointegrated Systems Made Easy", Working Papers in Econometrics and Applied Statistics No. 88, University of New England, 1996.

SACHS J., TORNELL A., and VELASCO A., "Financial Crises in Emerging Markets: The Lessons of 1995", *Brookings Papers on Economic Activity*, 1996, *1*, pp. 147-217.

SAID D.E. and DICKEY D.A., "Testing for Unit Root in Autoregressive Moving Average Models of Unknown Order", *Biometrika*, 1984, *71*, pp. 599-607.

SCHWERT G.W., "Effects of Model Specification on Tests for Unit Roots in Macroeconomic Data", *Journal of Monetary Economics*, 1987, *20*, pp. 73-103.

SHAN J., TIAN G.G., and SUN F., "The FDI-led Growth Hypothesis: Further Econometric Evidence from China", mimeo, Australian National University, 1997.

SMITH C., "Stock Markets and Exchange Rates: A Multi-Country Approach", *Journal of Macroeconomics*, 1992, *14*, pp. 607-29.

STOCK J.H., "Cointegration, Long-Run Comovements, and Long-Run Horizon Forecasting", mimeo. University of New South Wales, 1995.

TODA H.Y. and PHILLIPS P.C.B., "Vector Autoregressions and Causality", *Econometrica*, 1993, *61*, pp. 1367-93.

—— and YAMAMOTO T., "Statistical Inference in Vector Autoregression with Possibly Integrated Processes", *Journal of Econometrics*, 1995, *66*, pp. 225-50.

WEI K.C.J., YU-JANE L., CHAU-CHEN Y., and GUEY-SHIANG C., "Volatility and Price Change Spillover Effects Across the Developed and Emerging Market", *Pacific-Basin Finance Journal*, 1995, *3*, pp. 113-36.

ZAPATA H.O. and RAMBALDI A.N., "Monte Carlo Evidence on Cointegration and Causation", *Oxford Bulletin of Economics and Statistics*, 1997, *59*(2), pp. 285-98.

ABSTRACT

This paper investigates the causal behavior between the nominal Malaysian ringgit exchange rate (MYR/US) and the Kuala Lumpur main and second board stock indices. The results of the multivariate model show that the main board index leads the exchange rate during the crisis period, while the exchange rate leads the second board index during the pre-crisis period. In all other cases, there are bi-directional causal relationships between stock price indices and exchange rate. This indicates that the exchange rate could only be a useful predictive indicator for the second board index.

JEL classification: *G1, C5, F3*
Keywords: *exchange rate, stock prices, Malaysia*



RISEC, Volume 50 (2003), No. 4, 521-543

THE EARLY DEVELOPMENT OF BANKING IN ETHIOPIA

by
ARNALDO MAURI*

1. *Introduction*

The fundamental role of monetary and financial institutions for economic development is widely recognised by economic literature. Basic conditions, although rudimentary, in financial field conducive to economic advancement were made available in Ethiopia at the beginning of the XX century. In fact, the history of banking in this African country goes as far back as 1905 when the first bank was established.

There are five principal events, which may conveniently be taken as dividing Ethiopian banking history into periods. The first event was the establishment of the Bank of Abyssinia, marking the advent of banking into the country. The second event was the Italian occupation in 1936, when a broad colonial banking network, extended to encompass all Italian possessions in the Horn of Africa (Eritrea, Ethiopia and Somalia) and closely linked with the metropolitan financial system, was set up in the country (Mauri, 1967; De Rosa, 1982; Tuccimei, 1998). The third event was, in 1943, the creation of the State Bank of Ethiopia, marking the rebirth of the Ethiopian independent banking. This occurred during World War II after liberation of the country (Luther, 1958; Konczacki, 1962; Deguefe, 1973). The fourth event was the revolution of 1974, which wiped out the monarchy, nationalised companies and shaped a "socialist banking" two-tier modelled, the whole credit system being based on a central bank and three state-owned financial institutions, each of them enjoying monopoly in its respective

* Department of Economics and Business Administration, University of Milan, Milan (Italy). E-mail: arnaldo.mauri@unimi.it

market (Caselli and Mauri, 1986). The fifth event was the collapse of the socialist regime followed by a financial sector reform and liberalization according to the Monetary and Banking Proclamation of 1994, which brought out, as a tangible result, a broadening of the banking market and an increasing level of competition, these two trends being, obviously, interrelated.

This study concentrates on the first period of the banking history of Ethiopia (1905–1936) marked by the origins and early development of this industry in the country. The remainder of the paper is organised into five sections. The next two sections, 2 and 3, consider preliminary steps to the establishment of the Bank of Abyssinia, the first bank in Ethiopia, and reviews operations of the Bank in a backward context. Section 4 focuses on the entering of other financial institutions into the Ethiopian financial market and on the French interests involved. Section 5 describes the liquidation of the Bank of Abyssinia. Finally Section 6 reports on the establishment of the Bank of Ethiopia and evaluates the results achieved by this institution during its short life.

2. *The Project of Menelek II: Preliminary Steps*

Ethiopia may be listed among the longest surviving states in the whole world, its history going back to the ancient Axumite kingdom. Because of the mountainous nature of the territory and the indomitable spirit of its inhabitants, this inland country, unlike other African countries, enjoyed ages of political independence, interrupted only by a short period of Italian colonial rule. Ethiopia, in fact, has been able to preserve, more than any other country in the continent, its own national identity and cultural heritage. It is not a surprise, therefore, if one can hardly note in Ethiopia the distinguishing imprint of a single colonial power or foreign conqueror. This is true even in the financial sphere, where, given the backward conditions of the country at the beginning of XX century (Alamanni, 1891), it was necessary to have large recourse to foreign assistance and co-operation in setting up institutions and markets.

The Emperor Menelek II, as a second step after the institution of a national monetary system (Sellassié, 1932) wanted to establish a bank in the country, at that time called Abyssinia. For understandable motives of national pride he gave preference to an institution at least formally independent rather than to a much easier solution represented by opening a branch office of an expatriate bank. There are differing versions of the events, which led up to the

constitution of this bank. According to Pankhurst (1963), who relies primarily on British sources, the first contacts were made with France, but later Menelek's sympathies shifted towards Great Britain; therefore the important task of establishing a national bank in Abyssinia was entrusted only to that country.

A slightly different, and perhaps more complete version of the facts, suggests that the Emperor contacted envoys from a number of European countries asking to work together on this project (Rossetti, 1910; Mauri, 1967; Tuccimei, 1998). While political and financial aspects of the project were being discussed in France, in Italy and in other European continental countries, the prompt British answer reached Addis Ababa. London was ready to provide assistance in this venture and felt it was best to entrust the task to the National Bank of Egypt, a private bank, under British control registered in Egypt as a limited company. During negotiations, in 1905, in exchange for granting to the entering institution a banking and note issuing monopoly in Abyssinia lasting fifty years, Menelek put forward three conditions: (a) internationalisation of the institution on three tiers represented respectively by the shareholders, the Board of Directors and the management; (b) a number of seats in the Board to be reserved to Imperial nominees; (c) transfer to the Emperor of a substantial share of profits. The last clause might have been regarded as a consideration for the concession of monopoly, while the first two aimed at limiting British influence in the Abyssinian bank due to the fact that presumably there would have emerged conflicting positions of representatives of European countries within the Board of Directors. Drawing from the lesson of the railway project Djibouti-Addis Ababa (the concession granted by the Emperor to the French in 1894), according to Zewde (1991), a conscious effort was made, at that time, to internationalise the banking venture.

3. *The Bank of Abyssinia*

Negotiations got off well and eventually, on March 11th of the same year, a convention for establishing the bank was signed in Addis Ababa by the two parties (the Emperor of Ethiopia and the National Bank of Egypt). The convention, known as "Charter for the Bank of Abyssinia", drawn in a twin-language text (English and Amharic), had provisions which covered the main features to be conferred to the Bank and listed the following eight privileges guaranteed to this institution by the Ethiopian Government for 50 years: (1) banking monopoly; (2) the sole right to issue bank notes which were to be

legal tender in the country; (3) the sole right to mint coins; (4) to be the sole financial agent of the Government; (5) preference for all issues of Government loans; (6) permission to build bonded warehouses; (7) free grant of all the land needed for the Bank's premises; (8) transport of its specie by rail in the country at Government rates. The third article of the convention had a clause which Menelek felt very strongly about, requiring a 10% of profit to be placed in the reserve, a 7% dividend to be distributed to shareholders (presumably of paid-up capital), and 20% of the remainder to be assigned to the Emperor in compensation for the privileges granted to the Bank. The final article (Article IV) laid down that the concession granted by the Emperor would be null and void in the case of the Bank not properly constituted before the first of July 1906 (Mauri, 1967).

The Governments of Paris and Rome, being informed that something was hatching in Addis Ababa, approached the Foreign Office in London requesting clarification and claiming their right to participate in setting up the Abyssinian bank in the spirit of the Tripartite Treaty, still under discussion at that time (Schaefer, 1992). The British answer pointed out that, since the establishment of the first bank in Ethiopia was a purely private affair, French and Italian participation was not a matter to be discussed at government level; rather, there should be private negotiations between all the banks concerned and these talks should take place in Cairo at the head office of the National Bank of Egypt.

Meanwhile, in Cairo, the new bank had been established according to the Egyptian commercial law at the end of May 1905. The statute was drawn up bearing in mind all the clauses contained in the Convention of Addis Ababa. The share capital of the Bank of Abyssinia was set at 500,000 sterling pounds, divided into 5-pound shares. Of this authorized capital a 25% was to be paid-in at the time of the constitution. Cairo was chosen for the registered office of the institution and consequently as the forum for Board of Directors' and Shareholders' meetings. The statute called for the adoption of the pound sterling as the unit of account of the Bank of Abyssinia, even though most bank transactions would be carried out, as usual in Ethiopia, in Maria Theresa thalers (Petazzi, 1911); this feature was not to have unimportant consequences for the Bank management, since the money of account was on gold standard, while the specie normally used in Ethiopia was based on silver standard. The Governor of the National Bank of Egypt was to be the Board's ex-officio Chairman; he was to be a voting member as were the Governor of the Bank of Abyssinia (Vice Chairman of the Board) and an additional eight ordinary members. The latter were to be nominated by the Shareholders' meeting; at least three of them were to be chosen among the Board members

of the Egyptian bank and two among Ethiopian notables. No statutory provision applied to the remaining three seats on the Board of Directors. Thanks to the casting vote of the Board Chairman, the control of the bank was guaranteed to the British even in the unlikely event of a coalition among all other members, Europeans and Ethiopians.

On the British side, the promise to involve other countries in the project was renewed. In fact a meeting was set up between the National Bank of Egypt representatives and those of the French and Italian banking groups. At the meeting, held in Paris in June, 1905, most of the French and Italian requests were accepted and taken into consideration in drawing up the text of the syndication agreement, a compromise that left largely unaltered the governance of the Bank and, particularly, the British position of predominance. The 100,000 shares of the Bank of Abyssinia were assigned as follows:

- Anglo-Egyptian Group 50,000 shares
- French Group 25,000 shares,
- Italian Group 25,000 shares

However, all the shares were to be placed in a common account administered by the National Bank of Egypt, which had pledged to take up 29,000 shares for itself and for the eight directors; parallel pledges were made by the French and Italian groups for 13,000 shares each. In order to offer the remaining 45,000 shares to public subscription, a banking syndicate, located in Paris at the headquarters of the Société Générale de Crédit Industriel et Commercial, was formed, in which the British banking group held 50 per cent and French and Italian groups 25 per cent each. The Syndication agreement of July 17th, 1905, fixed the geographical distribution of share allotment for public subscription. The largest number of shares to be issued for public subscription was to be offered in the country where the Bank of Abyssinia was to operate, that is Ethiopia (7,500 shares). England, Egypt, France and Italy were to be treated equally (6,625 shares each). Slightly smaller was the lot to be placed in the German market (5,000 shares) and lastly smaller lots were to be allotted to the capital markets in the Austrian-Hungarian Empire and in the United States (3,000 shares each).

The Steering Committee of the Syndicate later decided the technical details of the allocation operation. The shares, whose face value was 5 pounds, one fourth of which paid-up, were offered to the public at a price of £ 3. 5 s. 0 d., since the Syndicate reserved for itself a premium of 2 pounds per share. The Bank of Abyssinia could request later the remaining payment. In addition subscribers were promised that the necessary steps would be taken to have the shares traded in the stock exchange. It was resolved to open the

public subscription for one day only, November 7th, through the following banks: National Bank of Egypt in London, Cairo and Alexandria, Société Générale de Crédit Industriel et Commercial in Paris, Banca d'Italia and Banco di Roma in Rome, Credito Italiano in Genoa, Banca Commerciale Italiana (Confalonieri, 1980) and Società Bancaria Italiana in Milan, Bank für Handel und Industrie in Berlin, Credit Anstalt für Handel und Gewerke in Vienna, Baring Magoun & Company in New York. Subscriptions were also collected in Ethiopia.

The placing of Bank of Abyssinia shares, as expected, had a considerable success. In fact 610,000 shares, i.e. more than thirteen times the total amount offered, were subscribed. At the conclusion, the 45,000 shares sold through public subscription were distributed by nationality as follows: 14,000 shares to Anglo-Egyptian subscribers, 7,000 shares to French subscribers, 7,000 shares to Italian subscribers, 10,820 shares to German, Austrian and American subscribers, 6,180 shares to Ethiopian subscribers. The final holdings of the 95,000 Bank of Abyssinia shares (set aside the 5,000 shares reserved to the Board's members) summing up distribution among members of the Syndicate and public subscription, were: Anglo-Egyptian holders 41.05% of the share capital, French holders 20.53%, Italian holders 20.53%, Ethiopian holders 6.50%, all other holders 11.39%.

In the first Board of Directors of the Bank of Abyssinia, besides the Chairman, the Governor of the National Bank of Egypt, three seats were assigned to directors of the Anglo-Egyptian bank, two to Ethiopian notables and three to the minority, one each to the French, the Italian and the German Groups (Mauri, 1967). The opening ceremony for the Addis Ababa head office of the Bank of Abyssinia was held on February 15th, 1906 in the presence of the Emperor and his retinue of court dignitaries and diplomatic envoys (Pollera, 1926). At the outset of its activities the Bank of Abyssinia ran into its first problems and obstacles; on the one side the Ethiopian Emperor, who perhaps too late had become fully aware of the British supremacy in running the Bank, and on the other side the Ethiopian public, yet unaccustomed to making use of bank services.

Menelek, first of all, found it difficult to accept that the most important administrative decisions of the Abyssinian bank were taken at the Board of Directors meetings held abroad, in Cairo, where the two Ethiopian directors were usually absent. He thought that in the negotiations preceding the signing of the Addis Ababa Convention, his incompetence had led him to ask for much less than would have been fair in the interest of his country considering the series of privileges conceded to the Bank's founders. He therefore felt that the Convention should be revised and, accordingly, presented a number

of requests dealing chiefly with the governance of the Bank (Mauri, 1967).

Perhaps the most important request of the Emperor concerned the formation of an Executive Committee in Addis Ababa, to be invested with wide-ranging delegation of powers from the Board of Directors in Cairo. The Executive Committee should be composed of three members: the Governor of the Bank of Abyssinia and the two Ethiopian directors. All the decisions of the Committee were to be unanimous. Other requests were also aimed at increasing the independence of the Bank in respect to the founding Egyptian bank and at the same time at making the Bank of Abyssinia, in the guise of a true State bank, dependent on the Ethiopian Government. In addition, the Emperor demanded that the Ethiopian authorities be given the right to take monetary decisions, such as fixing backing ratios for issuing notes, and banking regulation and supervision, such as the power to send at any time inspectors to the Bank. Menelek, while awaiting a reply from Cairo, adopted a non-cooperative attitude. In other words he did not make fully use of his authority to ensure that all pledges made in the Addis Ababa Convention were respected. Specifically point four of Article II, which appointed the Bank of Abyssinia exclusive financial agent of the Government was almost ignored by Ethiopian authorities. Furthermore Ethiopian Customs and Addis Ababa Post Office officials often refused payments made in notes issued by the Abyssinian bank (Rey, 1923). At the same time French and Italian hopes to increase the share of control of the Bank grew; to this end Paris and Rome put forward a proposal to set up a six-member Executive Committee located in Addis Ababa, with one French and one Italian representative (Mauri, 1967).

From the beginning of its activity the Bank of Abyssinia made conscious efforts to build up its administrative organization and lay out a network of branches in the country. Soon after the inauguration of the headquarters in Addis Ababa, a branch office was opened in Harar (Zervos, 1936), long known as a trading centre with the Somali coast. Two years later a branch was opened in Diredawa, a city that had become important after the completion of the first trunk of the Djibouti-Addis Ababa railway. In 1910 the Bank's new headquarters were inaugurated in Addis Ababa (Pankhurst, 1963) and in 1912 a branch was opened in Gore, with sub-offices in Gambela and, temporarily, in Dembidollo, all caravan centres for trade with Sudan. After an interval of about eight years embracing the entire World War I era, the Bank underwent a further network expansion: this was marked by the opening of a branch at Dessye (Wallo) on the caravan route linking Shewa to Eritrea and a foreign office in Djibouti.

The management of the Bank of Abyssinia from the very start of the

Bank's operations continued Menelek's cause of spreading the use of national currency throughout the country. Efforts were made to build up confidence on the exchangeability of the notes issued by the Bank (Teodorani, 1912). The widening of the monetary area of the Ethiopian economy was in fact a chief objective strongly pursued, although with differing ends, both by the Government and the Bank of Abyssinia: the former aimed at the country's progress, the latter at gaining the confidence of the public for the expansion of its business.

Table 1 shows, for the whole life of the Bank of Abyssinia, some selected items taken from its *Annual Reports*. Figures of Total Assets, Deposits and Dividends decided by shareholder's meetings of the institution are expressed in pounds whilst Outstanding Notes are in Ethiopian thalers. The peak of Total Assets as well as for Deposits was reached in 1919. The circulation of Bank of Abyssinia notes in Ethiopian thalers registered, instead, a sharp growth during the last three years of the Bank's life (1928-1930).

Table 1. Bank of Abyssinia

End of Year	Total assets *	Deposits *	Dividends *	Outstanding notes **
1907	184,848	16,951	-	-
1908	195,408	19,894	-	-
1909	209,851	27,461	-	-
1910	274,037	*n.a.*	-	-
1911	358,474	56,593	-	-
1912	435,651	*n.a.*	-	-
1913	391,396	81,393	-	-
1914	370,837	*n.a.*	-	-
1915	*n.a.*	*n.a.*	-	-
1916	416,673	*n.a.*	-	214,765
1917	590,228	*n.a.*	-	214,765
1918	747,744	*n.a.*	0.1.0	214,765
1919	1,005,540	480,699	0.1.3	214,765
1920	733,991	*n.a.*	0.0.9	214,765
1921	595,855	*n.a.*	-	214,765
1922	498,237	*n.a.*	-	214,765
1923	478,204	226,730	-	214,765
1924	502,562	228,317	0.1.0	214,765
1925	669,543	434,678	0.1.0	300,000
1926	541,283	305,344	0.1.3	300,000
1927	699,951	347,417	0.1.3	750,000
1928	684,466	338,202	0.1.3	1,324,000
1929	778,168	456,619	0.1.3	1,524,000
1930	539,710	252,139	0.1.3	1,740,000

(*) figures in pounds; (**) figures in Ethiopian thalers.
Source: Bank of Abyssinia, *Annual Reports*.

The content of the Addis Ababa Convention may lead one to think that the Bank could have full control of the money supply in Ethiopia, given that it minted coins for the State, issued notes under monopoly and was the only depository institution in the country. In truth the Bank's powers were greatly limited, since national specie and bank notes had a very narrow circulation, which did not reach, in most instances, beyond the major urban centres. At the same time, the Maria Theresa silver thaler, a full value commodity money, was in use everywhere not only in Ethiopia but also in the countries of the Red Sea Littoral (Santagata, 1935; Martini, 1939; Tschoegl, 2000). According to Pankhurst (1965) and other authors the Maria Theresa thaler, besides its use as a coin, i.e. means of payment, standard of value and store of value, also performed important non-monetary functions: in fact it served as a jewel, as a certified source of silver (content 833.33%) for melting down and even as a measure of weight (28.0668 grams). Coming back to the issue of notes by the Bank of Abyssinia, it has to be underlined that these notes had to be fully covered by silver coins. Lastly, there were no reserve requirements on deposits, but disappointedly for the Bank, they held only a marginal position among the kinds of money used for payments in the country. Bank notes and bank deposits, in fact, are to be regarded as interchangeable forms of bank liability for payments, but their use depends on the public's choices.

Because of the country's backwardness, which made transfer of money quite difficult and expensive, large cash balances in specie were needed in branches. It is not therefore too surprising to find out that earning assets represented less than half of the Bank's total assets. In part loans and investments were made in Egypt, primarily during the Bank's early years. The usual lending of the Bank of Abyssinia in Egypt was made by advances on securities, a type of transaction unknown in Ethiopia, where guaranteed loans, overdraft facilities and advances on goods, often stored in warehouses, were granted.

The annual interest rate charged on loans guaranteed by collateral was 12%. In general an advance equalled 70 per cent of the value of goods. Overdraft on current accounts were almost exclusively granted to companies run by foreigners and were much favoured by the Bank since they allowed for the expansion of the use of deposit money, a medium of exchange evidenced only by figures in books, which entailed creation of money. In fact, as mentioned before, no compulsory reserve requirements were applied on deposit balances.

From the first days of this venture a great part of bank credit was channelled to financing international trade. In particular, the Bank was involved on the one hand in financing the collection and export of the

primary commodities and on the other in financing distribution of imports of consumer goods. Its function was, of course, that of bridging the gap between the purchases of coffee, for example, from the farmers, to the time at which it was sold on international markets (Cora, 1913). Indigenous borrowers were discriminated by the adoption of "sound banking" canons tailored to suit needs and habits of more developed countries. For a limited period, however, mortgage loans were also made, but difficulties arisen in collection discouraged further transactions of this type. Finally are to be mentioned personal loans to Ethiopian citizens; these loans, requiring two guarantors, carried an interest rate of 15%. The higher rate on these transactions was due mainly to procedural differences in debt collection depending on the borrower's status, European or native. Ethiopian subjects were, in fact, governed by customary law, which was somewhat primitive. It should not come as a surprise, therefore, that the Bank of Abyssinia had set part of its premises to use as a prison where recalcitrant debtors were held until their relatives settled the debts, accrued monthly not only with interests, but also with expenses for maintenance of prisoners (National Bank of Egypt, 1948).

Investments in British Government bonds were also important as a second line liquidity reserve. During the World War, a large quantity of English National War 5% Bonds flowed into the Bank's portfolio. Even deposit operations were not as successful as had been hoped. In particular, the inflow of savings deposits was scarce because of a number of factors such as the modest branch network, the low level of per capita income, the early stage of monetary development of the country (very low monetisation ratio of the economy), the lack of confidence on the part of the public (Martini, 1939) and the unattractive interest rate offered to depositors, i.e. 3% per annum. Current account deposits recorded better levels, since they were used by foreign businessmen and rich Ethiopian clients more accustomed to banking services and to the use of drawing cheques. The Bank of Abyssinia also drew in considerable proceeds from exchange operations and imported huge quantities of Maria Theresa thalers, which were placed in the domestic market. Transfer of funds within the country was also a good source of revenue, since a 0.5% commission was charged for the service. International transfers however were even more profitable for the Bank.

A glance at Bank of Abyssinia series of Profit and Loss statements and an investigation of performance determinants reveals considerable variability of results due both to business climate changes (connected chiefly with Ethiopian exports) and fluctuations in the price of silver and, consequently, in the exchange rate Ethiopian thaler/pound sterling. Losses were recorded during the first years of operation but starting from 1909, profits are

recorded. However, caution dissuaded directors from paying dividends until 1918 (see Table 1) in order to recover preceding losses registered in the take-off stage and because of uncertainties during the war period. For two years after, dividends to shareholders were paid until a three-year slump of the Ethiopian economy (1921-1923). Payment of dividends recommenced in 1924 and continued without interruptions until 1930, when the Bank of Abyssinia went into liquidation.

At the end of 1930 the total amount of notes in circulation was 1,740,000 thalers (evaluated £. 88,549 12 s. 4.d in the Balance Sheet of Bank of Abyssinia) fully covered in the assets side by an equivalent amount of silver coins. Cash on hand in specie was almost £. 164 thousand and overall deposits more than £. 250 thousand.

4. *The Entry of Other Financial Institutions*

The Addis Ababa Convention of 1905 had guaranteed to the Bank of Abyssinia the country's banking monopoly for fifty years, but this pledge was not long maintained by Menelek. After just a few years, in fact, the Société Nationale d'Ethiopie pour le Développement de l'Agriculture et du Commerce (henceforth referred as SNEDAC) was founded. Unlike the Bank of Abyssinia, this bank was a wholly Ethiopian institution, even though the Emperor used the services of a Greek long time resident in Addis Ababa in this enterprise. With this exception, all personnel were Ethiopian as was the institution's share capital (Pollera, 1926; Pankhurst, 1963). One may ask first of all what objectives Menelek had in mind in making himself the promoter of bringing into being a second financial institution just a few years after conceding the banking monopoly to the Bank of Abyssinia; second, what might have motivated the lack of opposition by the Bank of Abyssinia to this new venture openly violating past agreements (Addis Ababa Convention of 1905).

As far as the first query is concerned, it must be recalled that the Emperor had watched disappointedly how the Bank of Abyssinia fell under British control through shareholding by the National Bank of Egypt and did not approve in particular the lending policy followed by the Bank. The Bank was blamed to be unable to find local employment for its funds. In fact this institution, as a result of the European background of its managers, followed certain standards of credit-worthiness and imposed certain conditions as to the security of loans to be disbursed. The credit policy of the Abyssinian bank, mainly oriented to financing foreign enterprises and wealthy clients in

the country and international trade on the one hand and rationing loans to domestic small borrowers as farmers, traders, craftsmen, and householders on the other appeared to be inappropriate. A policy of severe credit rationing was mainly implemented through non-price mechanisms of transaction costs allocation to the borrowers, i.e. practices of manipulating non-interest terms in lending operations. As a result, the Bank of Abyssinia loan portfolio was highly concentrated in favour of large non-rationed wealthy borrowers. Unsatisfied borrowers were forced to seek credit on informal financial market at much higher interest rates.

Menelek wanted therefore to establish a second bank in the country to fill the gaps of the Bank of Abyssinia lending. It was, as pointed out by Schaefer (1992), the Ethiopian response based on self reliance. The SNEDAC's loans, generally backed by personal guarantees, were granted at the annual rate of 10 per cent marking a substantial reduction over the 15 per cent charged by the Bank of Abyssinia on the indigenous loans. SNEDAC's lending was concentrated in a segment of the market for loans on purpose neglected by the Bank of Abyssinia and therefore the newly created financial institution was not felt as a competitor by the latter. In this way the conciliatory behaviour of the Abyssinian bank towards the newcomer may be easily understood.

Before reviewing the successive development of SNEDAC after French capital became involved, an historical digression must be made. The two major rival powers in Africa, the British and the French, had found a competition-ground also in Ethiopia. France, winner in the railway affair (Concession of 1894 for Djibouti-Addis Ababa railway), but looser – as well as Italy – in the banking affair (Concession of 1905 for the Bank of Abyssinia), did not renounce its aims of penetration into the Ethiopian financial market. The French strategy was however much more subtle and articulate than the Italian one, which was simply and exclusively focused on the governance of the Abyssinian bank. The French strategy ran along two main lines: while backing Italian claims for more equitable distribution of power at the Bank's summit, it tried at the same time to convince authorities in Addis Ababa to revoke unilaterally the 50-year privilege granted to the Bank of Abyssinia, so as to clear the way for the establishment of French financial institutions in Ethiopia. As hopes of increasing power within the Bank of Abyssinia faded, French efforts shifted gradually towards the second option.

Initially attempts were made to introduce financial services in the Abyssinian Post Office, which was a French domain; later the target of obtaining authorization for opening branch offices of the Banque de

l'Indochine prevailed. This bank, in fact, a few years earlier, had established a branch in Djibouti and claimed that its entry into the Ethiopian financial market would accelerate the completion of the work underway on the railroad linking Addis Ababa with the Indian Ocean port, main conduit of Ethiopia's foreign trade. Furthermore Djibouti office of Banque de l'Indochine had been successfully operating for some time with Ethiopian and expatriate merchants, mostly from Harar, offering better terms and conditions than those practiced by the Bank of Abyssinia (Ahooja, 1965). The Bank's forceful opposition however torpedoed a 1908 attempt of the French bank to open a branch in Diredawa, temporary terminus of the railway.

The Bank of Abyssinia different attitude towards other banks' attempts to enter the Ethiopian credit market is quite understandable. SNEDAC did not represent a real danger to the interests of the Abyssinian bank and did not pose a political threat to the British prominent position in the country. Instead, the acceptance of Banque de l'Indochine application for opening branches in the country by the Ethiopian authorities would have been dangerous from both an economic and a political point of view: Banque de l'Indochine would have been a strong competitor and its entry would have reinforced the French position in Ethiopia.

After the repeated attempts to open up Ethiopia to French banks had been made in vain, in the late 1920s, when it was clear that Addis Ababa Government had in mind to cancel the concession of the Bank of Abyssinia, Paris' attention focused on SNEDAC. A French financial group made the necessary contacts with Ethiopian authorities to hand over the control of this company. As a result, SNEDAC became the Société Nationale d'Ethiopie on August 1st, 1928; its capital was increased to one million thalers, of which 60% was subscribed by the Société Financière de France (Ouannou, 1961). The Société Nationale d'Ethiopie collected deposits, both savings and current accounts, and granted loans. Lending operations were: (a) mortgage loans to Ethiopian citizens owning real estate in Addis Ababa; (b) advances on import and export of goods; (c) discounting and overdraft facilities mainly to European business.

A further step on the way of French penetration into the Ethiopian financial market was made in 1934, when Banque de l'Indochine succeeded in opening a branch office in Addis Ababa. In the same years the Compagnie de l'Afrique Orientale obtained authorisation for opening an agency in the city (Pankhurst, 1963).

After the Italian invasion of Ethiopia and the constitution in 1936 of the empire of Italian East Africa, French bankers and traders chose to withdraw from the country. Consequently the Banque de l'Indochine branch office in

Addis Ababa was closed as well as the agency of the Compagnie de l'Afrique Orientale and the shares of the Société Nationale d'Ethiopie were sold to an Italian financial group. The last mentioned institution however continued operations under the new name of Società Nazionale d'Etiopia (Mauri, 1967).

5. *The Bank of Abyssinia Goes into Liquidation*

After Lij Yasu's overthrow, in 1916, Zawditu, whose sole qualification, according to Zewde (1991), was her birth as the daughter of Menelek II, was proclamed Empress and Tafari Makonnen was designated simultaneously regent of the empire and heir to the throne. Ras Makonnen had an impressive lineage and an extraordinary talent. He was open to foreign culture and very ambitious: since the beginning it was clear his climb to the throne. This goal was attained only after the death of Empress Zawditu, in 1930. The coronation of Ras Tafari Makonnen, under the style of Haile Selassie, brought, once again after Menelek II, an outstanding personality to rule Ethiopia.

The new Emperor continued the Menelek's policy aiming at independence, modernization and progress of the country. Haile Selassie, who had once been a member of the Board of Directors of Bank of Abyssinia, following this line, gave priority to the reform of the Ethiopian monetary and banking system. The managers of the Bank were blamed by public opinion to have preconceptions about the conduct of banking business, which were the result of their cultural background. These criticisms were not groundless, but it should be underlined that these foreign bankers brought in banking business good standards of integrity and incorruptibility. The people's criticisms did not influence, nevertheless, the Emperor's decisions on this matter and the managers of the Bank were not dismissed. In conclusion we can state today that the Bank of Abyssinia was of undeniable utility to the Ethiopian economy, but in the 1930's the views had changed and the Emperor thought that it was no longer possible for the Ethiopian Government to accept the fact that the only issuing bank operating in the country was owned and controlled by foreigners.

Haile Selassie made his decision promptly known to the National Bank of Egypt. The Egyptian bank was not particularly reluctant to withdraw from Ethiopia on the condition that adequate compensation would be paid. Thus negotiations began in Cairo for Bank of Abyssinia liquidation.

In retrospect, the legitimate and long standing Ethiopian aspiration of creating a truly national bank of issue did not necessarily imply the

liquidation of the existing Bank of Abyssinia. It would have been sufficient to cancel the Abyssinian bank's privileges of issuing currency, of banking monopoly and of exclusive right to act as fiscal agent for the Government. At the same time, a new Ethiopian state bank could have been created. Nevertheless, both parties' interests lay in liquidating the Bank of Abyssinia. On the one hand the National Bank of Egypt was not so much interested to remain in the Horn of Africa and continue doing business under the less favourable conditions brought about by the change in the political setting and, on the other hand, the Ethiopian Government strongly wanted to make use of the existing organization (assets, liabilities, personnel and premises) of the Bank of Abyssinia in order to speed up implementation of the project for a State bank. This solution of liquidation, in fact, would make possible to establish the new bank without delay.

Negotiations were concluded on April 16th, 1930 with an agreement calling for (Zervos, 1936):

a) revocation of concession and privileges granted to the Bank of Abyssinia in the 1905 Convention;
b) winding up of the Egyptian joint-stock company "Bank of Abyssinia";
c) liquidation of the Bank of Abyssinia through conveyance en bloc of its assets and its liabilities to the Ethiopian Government;
d) payment by the Ethiopian Government of a sum of 40,000 pounds to Bank of Abyssinia shareholders as a compensation for terminating the concession nearly 25 years before it was scheduled to end and for the Bank's goodwill.

In accordance with agreed conditions, the Ethiopian Government paid up to shareholders a total of £ 203,807. 7 s. 6 d. in instalments until July 1931. This amount was calculated by summing up the adjusted book value of the Bank of Abyssinia, the above mentioned compensation of 40,000 pounds for terminating in advance the concession and the interest (at a rate of 6% per annum) on the debt of the Ethiopian Government from the end of 1930 to July 16th, 1931, the day on which the last instalment was paid. An important agreement was signed with the Ethiopian authorities on July 13th, 1932 regarding the notes issued by the Bank of Abyssinia. The liquidators transferred to the new Ethiopian bank the specie reserves in Maria Theresa silver thalers, which covered issues of notes by the Abyssinian bank depositing them in a specific account entitled "Trustee for Redemption of Bank of Abyssinia in voluntary liquidation notes". The new bank agreed to provide the conversion service for these notes free of charge (possibly by means of replacement with its own notes) and to make, as a guarantee, a deposit in South Australia 5% bonds with an overall face value of 5,000

pounds with the National Bank of Egypt. The guarantee deposit was to be freed gradually as Bank of Abyssinia notes were redeemed by the new Ethiopian bank and destroyed under the control of the Egyptian consulate in Addis Ababa.

6. *The Bank of Ethiopia*

The Imperial Decree issued on August 29th, 1931 chartered the new bank as Bank of Ethiopia and also represented the first banking law ever passed in the country (Pellegrineschi, 1936). With this decree, in fact, a regime of ex-ante authorization for banking industry was established and this provision was justified by attributing the qualification of public interest to banking operations.

The Bank of Ethiopia was set up as a joint-stock company, with an authorised capital of £ 750,000 divided into 30,000 shares of 25 pounds each. Private citizens were supposed to become shareholders since two different needs had to be met. On the one hand, given the important functions assigned to the Bank of Ethiopia, the Government was to have full control over the Bank's management. On the other hand private contribution to the share capital could have made the Government contribution less burdensome. The two objectives were not irreconcilable and it was deemed that the best solution would be to differentiate shares into two classes. Class A shares, 18,000 in number, would have been meant for the Government, while class B shares, 12,000 in number, would have been placed by public subscription. The two above mentioned classes would have enjoyed equal rights, the only relevant difference lying in the appointment of the Bank's directors. The holders of the two classes of shares would have met separately: holders of class A shares (i.e. Government) to appoint six members of the Board and holders of class B shares to appoint the remaining four members. This formula would have ensured the Government's control over the Bank even with a minority stake in the equity capital. At the time of constitution, only half of the share capital, however, was paid up. At the closing of subscriptions, a total of 18,877 shares of the 30,000 offered had been subscribed. The results did not come up to expectations as the public showed no interest in the operation. In particular, the subscription failed to attract foreign capital that had greatly contributed to the success of the Bank of Abyssinia stock subscription. The Ethiopian Government held 18,003 shares, nearly 95.4 per cent of the company capital, and another 874 shares were in the hands of the public. The paid-up capital of the Bank of Ethiopia turned

out to be £ 235,692 10 s. 0 d. instead of the expected £ 375,000. The sum paid by the Government included the purchase price paid to the shareholders of the Bank of Abyssinia because of its contribution in kind by conferring the assets and liabilities of the liquidated bank.

According to the Imperial Decree, the Bank of Ethiopia notes were to be fully covered by reserves including: minted and ingot gold, foreign currencies (provided they were payable in gold), bank bills and acceptances in gold currency, Ethiopian and foreign short-dated government bonds and other assets of a similar nature. The various assets backing notes were not however perfectly fungible and ceilings had been established for some types of reserve-assets. The Bank of Ethiopia was made the depository of public funds and was to provide other normal banking services, as fiscal agent, to the Government.

The Bank began operations in November 1931 with the same Governor (C.S. Collier) appointed chief executive, and took over management, staff and premises of the ceased Bank of Abyssinia (Deguefe, 1965). Since only the name of the institution had changed, most of the public were unaware of the event, which nevertheless represented a very important step on the way toward the county's emancipation (Pankhurst, 1963). The Bank of Ethiopia however came into being with a more favourable position than the Bank of Abyssinia. Relations with Ethiopian authorities were in fact much smoother than they had been for the former bank since Bank of Ethiopia's Board members were appointed by the Government itself.

The Board of Directors' powers were wide-ranging, encompassing also the main aspects of monetary management in the country. The Board was entrusted with fixing borrowing and lending rates applied by the Bank. Bank notes coverage in specie was only partial and not full as for the Bank of Abyssinia. Eventually the Bank of Ethiopia could have been considered a central bank, since, besides enjoying the sole right of note issue and acting as financial agent of the Government, it was providing refinancing facilities to the Société Nationale d'Ethiopie. Therefore the Bank of Ethiopia may be placed undoubtedly among the first indigenous central banks established in Africa (Ahooja, 1965; De Kock, 1974).

The Bank of Ethiopia in its ability to influence the quantity of money and credit could foster the economic development of the country. The Bank however was run conservatively because the management, taken over from the Bank of Abyssinia, preferred to keep on with the customary cautious policy of the previous bank. The pound sterling was maintained as unit of account even though in transactions with customers the use of Maria Theresa thalers prevailed by far. The branch network was extended with the opening

of an office in Gojam, at Debra Tabor.

Efforts to spread the use of the national currency throughout the country were eventually unsuccessful, but it should be mentioned that the main causes, which prevented the achievement of this ambitious goal were external and out of control of the Bank. The Ethiopian Government decided to replace the silver standard of the Bank of Abyssinia (full coverage of notes circulation in silver coins) with a gold exchange standard system, based on the British pound sterling, even though only silver coins were circulating in the country. Frequent and sharp fluctuations in the value of silver (high volatility of prices in the past experience) had suggested to abandon the silver standard and to adopt a monetary system based on gold; therefore minted silver was to be allowed to continue circulating, but only as token money. There was no circulation of minted gold and the new Ethiopian golden thaler was in fact to be only a virtual monetary unit.

A little less than a month after the Imperial decree was promulgated, an international crisis of the gold standard occurred. The Bank of Ethiopia had to adapt to the new situation while its notes were covered by minted silver (Maria Theresa thalers) and foreign currency. Maria Theresa thalers served to maintain internal convertibility and thus to permit continued contacts with a domestic economy dominated by the Austrian silver coins as well as the settlement of trade balances with a number of neighbouring countries. Conditioned by its political and economic international ties, Ethiopia sought to remain linked to the British sterling even after this currency interrupted convertibility into gold on September 21st, 1931. The dual reserve system of the Bank of Ethiopia, due to the desire to link simultaneously the Ethiopian monetary unit to the pound sterling and to silver, while silver prices underwent considerable fluctuations, was a cause of serious problems for the Bank. In facing these disturbances the Ethiopian bank tried to stabilize the silver thaler/pound sterling exchange rate. In other words efforts of the Bank of Ethiopia were directed to neutralize the effects on the domestic market of fluctuations of silver prices on the London market. This policy caused a growth of smuggling on the borders of the country, mainly during periods marked by maximum divergence between the market price of silver and the Bank's official price (Bertone, 1936).

Despite the difficulties, the Bank of Ethiopia pursued the objective of increasing the use of the national currency in the country and to enlarge its sphere of influence on the economy. Another objective of the Bank's monetary policy was to avoid disturbances due to seasonal movements. Attempts were made in the first place to safeguard what little progress had been achieved by the Bank of Abyssinia. The new bank notes therefore had

format, languages and appearance similar to that of the previous ones. Issue of notes, under the new legislation, began in May 1932 in the following denominations: 5, 10, 50, 100 and 500 Eth. thalers. In June 1933 notes in 2 Eth. thalers denomination were also issued (Zervos, 1936). As the new notes were issued Bank of Abyssinia old notes were withdrawn from circulation. At the end of 1933 nearly 94% of the outstanding circulation of Bank of Abyssinia notes had been redeemed. Ethiopian monetary authorities had in mind to establish confidence in the new currency both at home and abroad. While the note substitution went on smoothly, it proved a much harder job to expand the circulation of paper money to other segments of the domestic economy and attempts to replace all around the country the Maria Theresa thaler by the national currency proved unsuccessful (Shaw, 1935). Curiously enough, in some instances Bank of Ethiopia notes seem to have been accepted only at discounted value (Farago, 1935). Table 2 shows the circulation of notes issued by the Bank of Ethiopia; figures refer to the end of the month. The table indicates that the overall circulation of Bank of Abyssinia notes had been surpassed in October 1932, but reveals a standstill the following year. The peak was reached in February 1935 and from the following month began to shrink owing to impending war. The table enlightens another interesting phenomenon still visible in Ethiopia today: the seasonal fluctuations of money supply due to exports which usually make an inflow of foreign currency during the first quarter of the year, creating monetary base.

Table 2. Oustanding notes of Bank of Ethiopia

End-of-month circulation (unit: Eth. thalers)				
Month	Year			
	1932	1933	1934	1935
January	-	2,282,680	2,336,531	2,989,977
February	-	2,459,305	2,541,601	3,143,080
March	-	2,519,270	2,530,424	2,374,173
April	-	2,358,320	2,625,174	2,709,780
May	123,905	2,148,920	2,708,029	2,401,797
June	976,580	2,127,145	2,728,366	2,306,276
July	1,287,395	2,108,520	2,800,274	2,439,442
August	1,327,155	2,194,560	2,785,261	2,001,235
September	1,616,425	2,155,630	2,859,505	1,925,300
October	1,858,500	2,209,479	2,591,692	1,658,243
November	1,839,305	2,288,418	2,791,465	1,741,746
December	2,065,135	2,121,632	2,890,620	1,804,322

Source: Bank of Ethiopia, *Annual Reports.*

In order to correctly evaluate the results achieved by the Bank of Ethiopia in developing bank notes circulation in the country it is necessary to take into consideration the circulation of foreign-minted silver coins during the same period of time. At this point it should be added that the monetary reform also involved token coins by establishing a decimal based fractional currency system of nickel and bronze coins, fully backed by Maria Theresa thalers in the Bank's vaults.

The official stock of money within the country was, at that time, slightly less than 6.4 million thalers made up approximately as follows: bank notes 44%, token coins 11%, bank checking deposits 45%. Alongside this official money supply, circulated theoretically a mass of full value commodity money (Maria Theresa thalers) estimated at 40 - 50 million pieces (Mauri, 1997).

The Bank of Ethiopia co-existed with other financial institutions (see Section 4), each catering to the needs of different types of customers and with little overlap of functions. The Bank was engaged in collecting deposits in domestic currency as well as in a number of foreign currencies. There were no provisions for reserve requirements in respect to this liability similarly as it was with the Bank of Abyssinia. Demand deposits, chiefly made up with public funds, were not interest bearing and savings deposits accounted for one fifth of total deposits. Savings passbooks earned a modest rate, which did not reflect market conditions. This fact, together with a lack of a savings promotion campaigns, explains the unsatisfactory growth rate of deposit balances.

Earning assets were limited by the need to keep high idle cash balances in branch offices. Investment portfolio held Ethiopian and foreign Government bonds and bills. In addition the above mentioned South Australia 5% bonds were deposited as security for notes redemption with the National Bank of Egypt. In lending operations a very important role was represented by discounting of international and national bills and promissory notes. Advances on securities and merchandise, whose term almost never exceeded six months, were also granted. Merchandise was stored in the Bank's warehouses while precious metals were held in the Bank's vaults. An interest rate of 8-10% was applied to these transactions. In addition, the Bank of Ethiopia continued to grant loans to Ethiopian individuals covered by personal guarantees. Mortgage loans developed at a very modest rate. In conclusion, it is undeniable that lending by the Ethiopian bank during the first years of operations was excessively prudent and that this institution did not take full advantage of its leading position to implement a lending policy conducive to speeding up economic development (Balogh, 1964). However, the judgement should not be too severe because of the short life of this

institution. In fact the Italian invasion of Ethiopia interrupted the take off of this bank.

The Italian authorities, after the fall of Addis Ababa on May 5th, 1936, initially allowed normal activities of the Bank of Ethiopia to be continued because some time was needed to make decisions on the Bank's future. They were faced in fact with two alternatives equally viable, but each tied to different views on the target structure of the banking system of the Horn of Africa. On the one hand there was the opportunity of keeping the Bank alive, changing its ownership (class A registered shares being transferred from the Ethiopian Government to the Italian Government), name, management and, partially, staff. This line of action, which had already been successfully followed dealing with the Société Nationale d'Ethiopie, nonetheless entailed the adoption of an autonomous stance in the monetary and credit field which had not been taken in the past in the former colonies of Eritrea and Somalia, but which did not lack prominent supporters among Italian economists (Griziotti, 1937; Cosciani, 1939; Fanno, 1952). On the other hand the view calling for a complete integration of Ethiopia as well as other Italian possessions in East Africa into the Italian monetary and banking system necessarily implied the liquidation of the Bank of Ethiopia. The second alternative, mainly based on political issues, finally prevailed.

REFERENCES

AHOOJA K., "The Bank of Abyssinia: A Portrait", *Ethiopia Observer*, No. 4, 1965, pp. 355-57.

ALAMANNI E.Q.M., *La colonia Eritrea e i suoi commerci*, Torino: Bocca, 1891.

BALOGH T., *The Economic Impact of Monetary and Commercial Institutions of a European Origin in Africa*. Cairo: National Bank of Egypt, 1964.

BANK OF ABYSSINIA, Annual Reports, various years.

BANK OF ETHIOPIA, Annual Reports, various years.

BERTONE M., "Il problema monetario dell'Etiopia italiana", *Il Sole*, 25th and 26th July 1936.

CASELLI C. and MAURI A., *Moneta e banca in Etiopia*, Milano: Giuffrè, 1986.

CONFALONIERI A., *Banca e industria in Italia: 1894-1906*, vol. II, *Il sistema bancario tra le due crisi*, Bologna: Il Mulino, 1980.

CORA G., *Il commercio di esportazione dell'Etiopia nel 1911*, Roma: Ministero delle Colonie, Rapporti e monografie coloniali, 1913.

COSCIANI C., "L'organizzazione creditizia coloniale", *Rassegna Economica dell'Africa Italiana*, January 1939.

DEGUEFE T., "Credit Institutions in Ethiopia: An Historical Outline", *Ethiopia Observer*, No. 4, 1965, pp. 324-30.

——, *A Guide to Service*, Addis Ababa: Commercial Bank of Ethiopia, 1973.

DE KOCK M.H., *Central Banking*, London: Crosby Lockwood Staples, 1974.

DE ROSA L., *Storia del Banco di Roma*, 3 voll., Roma: Banco di Roma, 1982.

FANNO M., *La teoria economica della colonizzazione*, Torino: Einaudi, 1952.

FARAGO L., *Abyssinia on the Eve*, London: Putnam, 1935.

GRIZIOTTI B., *Il problema finanziario per la valorizzazione dell'Africa*, Sezione lombarda dell'Istituto Coloniale. Centro Studi, No. 4, 1937.

KONCZACKI Z.A., "Currency and Banking in Ethiopia", *The South African Journal of Economics*, September 1962, *30*(3), pp. 205-20.

LUTHER E., *Ethiopia Today*, Stanford: Stanford University Press, 1958.

MARTINI F., *Il diario eritreo*, Firenze: Vallecchi, 1939.

MAURI A., *Il mercato del credito in Etiopia*, Milano: Giuffré, 1967.

——, "The Rise of Banking in Ethiopia", *African Review of Money, Finance and Banking*, No. 1-2, 1997, pp. 39-56.

NATIONAL BANK OF EGYPT, *National Bank of Egypt: 1898 – 1948*, Cairo, 1948.

OUANNOU J., "L'Ethiopie et son économie", *Revue de la Société d'Etudes et d'Expansion*, November-December 1961.

—— and OUANNOU J., *L'Ethiopie pilote de l'Afrique*, Paris : Maisonneuve & Larose, 1962.

PANKHURST R., "Ethiopian Monetary and Banking Innovations in the Nineteenth and Early Twentieth Centuries'', *Journal of Ethiopian Studies*, 1963, *1-2*, pp. 4-120.

——, "The History of Currency and Banking in Ethiopia and the Horn of Africa from the Middle Ages to 1935", *Ethiopia Observer*, No. 4, 1965, pp. 358-408.

PELLEGRINESCHI A.V., *Etiopia, aspetti economici*, Milano: Principato, 1936.

PETAZZI E., "Il problema monetario dell'Eritrea", *Rivista d'Africa*, November 1911.

POLLERA A., *La vita commerciale etiopica e la circolazione monetaria eritrea*, Tivoli: Istituto Coloniale Italiano, 1926.

REY C., *Unconquered Abyssinia*, London: Seley Service & Co., 1923.

ROSSETTI C., *Storia diplomatica d'Etiopia durante il regno di Menelik II*, Torino: Società Tipografico-editrice Nazionale, 1910.

——, *Il regime monetario delle colonie italiane*, Roma: Loescher, 1914.

SANTAGATA F., *La Colonia Eritrea nel Mar Rosso davanti all'Abissinia*, Napoli: Libreria Internazionale Treves, 1935.

SCHAEFER C., "The Politics of Banking: The Bank of Abyssinia, 1905-1931", *International*

Journal of African Historical Studies, 1992, *25*(2).

SELLASSIE G., *Chronique du Règne de Ménélik II, Roy des Roys d'Ethiopie*, Paris : Librairie Orientale et Américaine G. Maisonneuve, 1932.

SHAW J.H., *Ethiopia*, New York, 1935.

TEODORANI P., *Commerci e dogane nel Tigrai*, Asmara: Tipografica Francescana, 1912.

TSCHOEGL A.E., "Maria Theresa's Thaler: A Case of International Money", *Eastern Economic Journal*, 2001, *27*(4).

TUCCIMEI E., *La Banca d'Italia in Africa*, Collana Storica della Banca d'Italia, Bari: Laterza, 1998.

ZERVOS A., *L'Empire d'Ethiopie: le miroir de l'Ethiopie moderne*, copyrighted in Athens and printed in Alexandria, 1936.

ZEWDE B., *A History of Modern Ethiopia 1855-1974*, London: James Currey, Athens, OH: Ohio University Press, Addis Ababa: Addis Ababa University Press, 1991.

ABSTRACT

This paper highlights the development of Ethiopian banking from the beginning of the XX century to 1936. In fact the history of banking in Ethiopia dates back to the turn of the century, when the Bank of Abyssinia was established. The National Bank of Egypt having been entrusted of the project, the new institution was chartered in Cairo and its shares were subscribed in a number of countries besides Ethiopia. The Bank of Abyssinia was given a 50-year concession and was engaged in issuing notes, collecting deposits and granting loans. Haile Selassie, after acceding to the throne in 1930, decided the liquidation of the Bank of Abyssinia. A new bank, the Bank of Ethiopia. under the Government control, was established in 1931 and retained management, staff, premises and clients of the old bank. Italian occupation in 1936 brought the liquidation of the Bank.

JEL classification: *O16, N27*
Keywords: *African banking, Ethiopia, Bank of Abyssinia, financial development*

RISEC, Volume 50 (2003), No. 4, 545-568

MANAGEMENT QUALITY, CREDIT POLICY, AND LOAN-LOSS DIFFERENCES AMONG COMMERCIAL BANKS IN NIGERIA

by
RALPH I. UDEGBUNAM*

1. *Introduction*

Financial distress and dramatic loan losses associated with recent financial deregulation were the proximate causes of bank failures in Nigeria in the 1990s. Deregulation fostered a highly competitive financial environment, as it opened the door wide for an unusually large number of new banks and non-bank financial institutions to enter the financial services industry. With the stiff competition banks' profitability, as measured by return on assets, came under severe pressure and plummeted, in many cases, to a very low level. To bolster the sharply declining profitability, most of these banks tended to adopt financial strategies that encouraged excessive risk taking. This is evident in the recent acute loan quality problems and large loan losses among commercial banks.

Apart from intense competition within and from outside the banking industry, banks were also faced with persistent economic downturn, virulent inflation, heightened political instability, escalating incidences of fraud and defalcation, and worsening economic and financial conditions of their corporate customers (see Udegbunam, 2000). The combined effects of all these factors are increased number of delinquent borrowers, phenomenal rise in non-performing loans and loan losses, sharply declining average

* Department of Economics and Statistics, Faculty of Social Sciences, University of Benin, Benin City (Nigeria). E-mail: rudegbunam2002@yahoo.com

I would like to thank an anonymous referee for his helpful comments and suggestions.

profitability, rapid capital erosion, and bank failures. But although the banking problems were widespread, the severity of loan problems varied greatly across the banks. Indeed, while some banks collapsed under the pressure of mounting loan-losses, others appeared to be recording high loan portfolio performance and increasing profitability. The question is: what can account for this disparity in the banks' loan portfolio performance? Can it be attributed to external factors beyond the control of the bank managers, or is it better explained by the banks' internal factors (or unique characteristics), such as credit policy and management efficiency?

This study uses a sample of 24 banks to examine the role of management quality and credit policy in the recent surge in loan problems at the Nigerian banking industry. The deterioration of loan quality and the escalation of loan losses can be explained by both external factors, such as economic, regulatory, and political conditions, and bank internal factors. Recent studies by Keeton and Morris (1987), Berger and Deyoung (1997), and Jordan (1998a), show that while a substantial part of the variations in problem loans and loan losses is due to differences in local economic conditions, a greater part of it is explained by differences in bank-specific attributes. However, a similar study by Huh and Kim (1994) for Japan and Korea finds evidence that is not totally consistent with the findings of the above studies that used a sample of United States commercial banks, as it indicates the dominant role of internal factors.

These results may not be surprising, as they reflect the structure of the banking system. United States has a predominantly unit banking system that usually restricts banks to operate in a particular locality often with a single office. In such a system, the local economic condition is an important determinant of bank performance. In contrast, Nigeria has a predominantly branch banking system that does not restrict, but permits banks to have a network of branches often scattered all over the country. In this system, the local economic condition and other external factors are not expected to exert a significant differential effect on the banks' financial performance.

The purpose of this study is to examine the role of bank-specific attributes, in particular credit policy and management efficiency, in the disparity in loan problems across the commercial banks in Nigeria. This will be beneficial not only to bank managers, but also to the regulatory and supervisory authorities. The remainder of the paper is organised as follows: Section 2 provides a theoretical basis for differences in loan losses across the commercial banks in Nigeria. In Section 3, the empirical methodology and model are presented along with the description of data used. The empirical results are presented and discussed in Section 4. Section 5 suggests some policy implications

while Section 6 concludes the paper.

2. *Theoretical Basis for Loan Loss Diversity*

To provide a background for the empirical analysis, we examine some theoretical reasons why there is diversity in the severity of the recent surge in loan problems among commercial banks. A variety of factors, some internal and some external, are responsible for the variation in loan problems and loan losses experienced by commercial banks in Nigeria (Udegbunam, 2000). Internal or bank-specific factors are those factors that are directly related to the individual bank operating strategies. They reflect the bank's overall credit policy, management quality, capital adequacy, leverage, risk preferences, asset growth, liquidity, culture and others. This set of factors is under the control of the bank management. External factors, which include economic, social and political conditions, regulatory and policy environment, and market structure, are outside the direct control of the bank management.

Although the focus of this study is on bank-specific or internal factors, in particular management quality and credit policy, it may be necessary to put the study in full perspective by briefly examining the role of external factors. Undoubtedly, political instability and weakened national economy are at the centre of the crises in the Nigerian banking industry. A number of studies including Keeton and Morris (1987), Emmons (1993) and Randall (1993) strongly suggest that banks are highly vulnerable to downturns in national, regional, and local economic conditions. With persistent political instability, social upheaval, and weakened national economy, Nigerian commercial banks faced an increasing number of delinquent borrowers and mounting problem loans. In the prevailing conditions, most of the businesses that constituted the banks' loan customers could not service their loans, and many defaulted in loan repayment.

The recent financial reform has also played a number of important roles in credit quality problems and bank failures in Nigeria. First, financial liberalization and the consequent explosion in the number of banks and nonbank financial institutions have greatly intensified competition in the banking industry. This had an adverse effect on bank profitability; and with sharply declining profitability, bank managers became increasingly motivated to take excessive credit risks. The result was large loan losses, insolvency and bank failures. Second, in a regulated financial system, lending rates are either pegged or are under ceilings. This does not encourage managers to make loans to high-risk customers, as it is not profitable. But with deregulation,

ceilings are removed and banks are motivated to make riskier loans, which promise greater return (see Demirguc-Kunt and Detragiache, 1997). The resultant acceleration in loan growth was found to cause a surge in loan losses (see Keeton, 1999; Sachs et al., 1996). Third, it is also pointed out by Demirguc-Kunt and Detragiache (1997), that bank managers trained in a highly regulated financial system may find it extremely difficult to cope with complex risk management and loan monitoring in a deregulated financial environment.

Since banks in Nigeria are almost equally exposed to external factors, the major causes of the diversity in the severity of loan problems are banks specific factors. Although we are primarily concerned with the role of management quality and credit policy, diversity in loan losses is also related to other important bank attributes. As control variables, it may be very useful to offer some theoretical reasons why they cause disparity in loan losses among commercial banks.

Management inefficiency could be the cause of the recent astronomical rise in the loan losses of the Nigerian commercial banks. A large body of theoretical and empirical studies appears to suggest that success or failure of a bank depends on the quality of its management (see Pantalone and Platt, 1987; Iyoha and Udegbunam, 1998; Wheelock and Wilson, 1995). Bank managers make a myriad of decisions in respect of allocation of bank resources, internal control, operating expenses, strategic planning, and loan policies. The surge in the number of new banks since deregulation and consequent sharp increase in the demand for trained bankers resulted in high labour turnover, acute shortage of skilled bankers, and employment of inexperienced managers by banks. Thus, the poor management especially of the new banks could explain the high rate of problem loans and loan losses of these banks.

Efforts requiring banks to maintain adequate capital have a long history (see Mitchell, 1909). Adequate capital has long been recognised as an important requirement for safety and soundness of a bank (see Gilbert et al., 1985; Gilbert, 1991; Estrella, 1995). Undoubtedly, inadequate capital increases the probability of bank failure; capital provides for asset expansion, absorbs losses, provides cushion against risk, and maintains public confidence in a bank. Recent studies by Peek and Rosengren (1996) and Estrella et al. (2000) suggest that capital ratios are the early warning indicators of banks' problems. The recent deterioration of the banks asset quality has had adverse effects on capital and earnings. Most of these banks have become grossly undercapitalised, and with sharply falling profitability, the managers of these banks have become increasingly inclined to take

excessive risks. Published studies have established that highly undercapitalised banks have increased incentive to make loans that have high probability of default (see Gilbert, 1991; Furlong, 1992; Edwards and Mishkin, 1995).

High financial leverage is closely related to high capital risk; it exposes a bank to large and mandatory interest payments – a significant claim on the banks' cash flow that must be met even when the bank is encountering losses. Undercapitalised banks tend to have greater incentive towards high levels of financial leverage (Koch, 1992, pp. 739-762). This suggests that the managers of banks with a high level of financial leverage would be more disposed towards risky loans and other high risk financial activities.

A bank's credit policy is an important determinant of its long-term financial performance. The structure and quality of a bank's loan portfolio largely reflect the bank's loan policy. Poor credit policy usually reflects poor management quality; it may also be deliberate. In a highly competitive environment, banks may deliberately reduce their credit standards in order to bolster declining profitability. However, poor credit policy, whether it is deliberate or is due to poor management, is expected to be closely associated with high credit risk, loan losses, and poor performance.

3. *Empirical Methodology and Data*

This section attempts to specify a simple model relating problem loans and loan losses to bank internal factors or specific attributes. We focus on bank internal factors because, since with a branch banking system, banks operate in the same economic, financial, political, policy, and regulatory environment, they are exposed to the same external factors. The major objective of the study is to examine the role of management inefficiency and poor credit policy in the recent astronomical rise in problem loans and loan losses of the Nigerian commercial banks.

Following recent studies by Huh and Kim (1994), and Jordan (1998a), a simple multiple regression model is specified to test for the influence of management quality and credit policy on banks' problem loan rate. However, unlike Huh and Kim and Jordan, this is essentially a cross-sectional study that focuses on the bank's management quality and credit policy. Other bank-specific attributes are included to control for their influences on cross-sectional differences in problem loan rates. Specifically, measures of capital adequacy, credit risk, and leverage are included as control variables. The simple regression model for problem loan rate is specified as:

$$(1) \qquad PLR_{it} = \alpha_{it} + \sum_{j=1}^{J}\sum_{k=1}^{K} \beta_{jk} X_{itk} + \varepsilon_{it} \qquad i = 1, \ldots, N, \; t = 1, \ldots, T$$

where PLR_{it} is the problem loan rate of bank i at time t, X_{itk} is a set of k bank i internal characteristics (independent variables) including management quality and credit policy at time t, β_{jk} is a set of parameters to be estimated, which measures the sensitivity of problem loan rate to bank-specific attributes X, α_{it} is the constant term and ε_{it} is the stochastic error term. Also N is the number of cross-sections, T is the length of time series, K is the number of independent variables, and j is the number of models to be estimated. Since the study covers a period of three years, 1995-1997 and with sample size of 24 banks, T is 3 and N is 24. From Table 1 there are eleven independent variables and thus K is 11, and for six models j is 6. An expanded form of equation (1) is the combined cross-section and time series of the problem loan rate model, which can be expressed in a more compact form as:

$$(2) \qquad PLR = \alpha + \beta X + e$$

where *PLR* is as defined, X is a vector of independent variables (bank-specific attributes), β is a vector of coefficients to be estimated, α an intercept term, and e is the error term. The independent variables that constitute the elements of the vector X are *BER, EQRA, LAR, LDR, UNSLL, LLPEQ, LATML, NITA, PTEA* and *EXLR* (see Table 1).

Variables of the model. - A variety of financial ratios, which are computed from publicly available sources, specifically balance sheet and income statement of the sample banks, are used as proxies for both dependent and independent variables. For the dependent variable we used the ratio of non-performing loans to total assets, which is denoted by *NPLTA*; that is, problem loan rate (*PLR*) is measured as non-performing loans-to-total assets ratio.

Non-performing loans are usually loans that are 90 days or more past due, or in nonaccrual status. However, in this study non-performing loan is used in a broader sense to include all bad-loans including charge offs. A high percentage of non-performing loans are ultimately bad and are charged off (see Gilbert et al., 1999).

The financial ratios (bank-specific attributes) included as independent variables or as components of X vector in equation (2), are categorized into five specific areas of bank performance: capital adequacy, leverage, credit policy, credit risk, and management quality. The variables under each

category are clearly defined in Table 1. We measured credit policy with four different financial ratios: the ratio of unsecured loans to total assets (*UNLTA*), the ratio of unsecured loans to total loans (*UNSLL*), the ratio of loan loss provision to equity capital (*EQRA*), and the ratio of loans above 12 months to total loans (*LATML*). Management quality is gauged with three different financial ratios: the ratio of net income to total assets (*NITA*), the ratio of profits to total assets (*PTEA*), and the ratio of expenditure to total liabilities (*EXLR*).

Table 1. Variable definitions

Variable	Definition	Expected sign
A. Problem loans measure		
NPLTA	Non-performing loans-to-total assets ratio	Dependent variable
B. Capital adequacy		
2. *EQRA*	Equity-to-risk assets ratio	–
C. Leverage		
3. *BER*	Borrowing-to-equity ratio	+
D. Credit policy		
4. *UNLTA*	Unsecured loans-to-total assets ratio	+
5. *UNSLL*	Unsecured loans-to-total loans ratio	+
6. *LLPEQ*	Loan loss provision-to-total loan ratio	+
7. *LATML*	Loans above 12 month-to-total loan ratio	+
E. Credit risk		
8. *LAR*	Loans-to-total assets ratio	+
9. *LDR*	Loans-to-total deposits ratio	+
F. Management quality		
10. *NITA*	Net income-to-total assets ratio	–
11. *PTEA*	Profit-to-earning assets ratio	–
12. *EXLR*	Expenditure-to-local liabilities ratio	+

In the case of control variables, we used the ratio of equity capital to total risk assets (*EQRA*) for capital adequacy. Financial leverage is measured as borrowings to equity capital ratio. Credit risk is measured with two financial ratios: the ratio of total loans to total assets (*LAR*), and the ratio of total loans to total deposits (*LDR*).

The expected signs of the coefficients of the independent variables (financial ratios), as indicated in Table 1, suggest that the severity of loan problems at the Nigerian commercial banks will be positively related to poor credit policy, poor management quality, credit risk, and leverage. We expect

that the indicator of the commercial bank's loan problems, non-performing loans/assets (*NPLTA*) will have a positive relationship with financial leverage measured by borrowings/equity (*BER*), all measures of credit policy (*UNLTA, UNSLL, LLPEA, LATML*), all measures of credit risk (*LAR, LDR*), and a measure of management quality (*EXLR*). The other two measures of management quality (*NITA* and *PTEA*), and the measure of capital adequacy (*EQRA*) are expected to be inversely related to the indicator of the commercial banks' loan problems (*NPLTA*).

All things being equal, lower levels of capital is expected to increase the likelihood of loan problems; banks with deficient capital are usually financially troubled banks that tend to have greater inclination for excessive risk taking. Loan problems are expected to be more severe in banks with high levels of financial leverage (*BER*), poor credit policy, high credit risk, and poor management quality. That is, banks with higher values of the ratios, *BER, UNLTA, UNSLL, LLPEA, LATM, LAR, LDR*, and *EXLR*, and lower values of financial ratios *EQRA, NITA*, and *PTEA* are expected to suffer severe loan problems. On the other hand, the more articulate and sound the credit policy, the greater the earnings and profitability, the better the quality of management and assets, and the better capitalized is a bank, the less is the likelihood of severe loan problems. The quality of management is crucial in credit policy, credit administration, and in the overall asset quality. Poor credit policy manifests in loan problems, the severity of which depends on the intensity of the other factors. Banks experiencing falling profitability and poor operating performance tend to have a high propensity to take greater credit risk in order to bolster falling profits. These banks often end up with severe loan problems and greater chances of failure.

Differences in individual bank loan problems and loan losses may also be related to bank size factors such as total assets and total deposits. In this study, as in most other similar studies (see Hunter and Srinivasan, 1990; Udegbunam, 2000), the use of total assets and total deposits as denominators in computing most of the financial ratios has considerably controlled for differences in bank size. For location, we have noted that banks in Nigeria are not restricted in their operations geographically, but are free to establish in any part of the country, and can have a network of branches. For these reasons, it is assumed that differences in loan problems across commercial banks, due to bank size and location, would not be significant. Thus, banks' internal characteristics are assumed to be the most important factor determining the differences in problem loans among commercial banks in Nigeria, and that of all these internal factors, credit policy and management quality are the most outstanding.

The data. - The sample consists of 24 commercial banks that reported non-performing loans according to the recently issued prudential guidelines. The guidelines issued in 1993 by the Central Bank of Nigeria (CBN), require banks to classify their loan portfolio into performing, non-performing, substandard, and bad.

The data used are from the sample banks' balance sheets and income statements. The data are essentially cross-sectional but pooled over three years, 1995-1997. In a pooled cross-section and time series regression, heteroskedasticity is always a problem. A frequently used method to detect it is that proposed by White (1980). However, in our estimation we assumed an error-component model. This is because the disturbance term (ε) in our model is likely to be composed of three independent components: cross-section error component; time-series error component, and combined time-series and cross-section error component[1]. In the model specified in equation (1) we have noted that N is the number of observations in the cross-section, T is the length of time series, and K the number of independent variables.

Following Fuller and Battese (1974), the appropriate estimation technique for error component models is a generalised least-squares method. This estimation method is used in our regression; it generates unbiased, efficient and consistent estimates.

4. *The Empirical Result*

The full results of estimating equation (1) by generalised least-squares (GLS) method, on pooled cross-section and time series data, are presented in Table A1 in Appendix A. Table 2 presents the preferred or the more parsimonious version of the results, which are reported in column 6 of Table A1.

---

[1] This involves some estimation issues. In the error components model, the stochastic terms ε_{it} in equation (1) are assumed tc have components as follows:

$$\varepsilon_{it} = u_i + v_t + w_{it}$$

where $u_i \sim N(0, \sigma_u^2)$ = Cross-section component
$v_t \sim N(0, \sigma_v^2)$ = Time series component
$w_{it} \sim N(0, \sigma_w^2)$ = Cross-section/time series component

The individual error components are assumed to be uncorrelated with each other and not to be auto correlated. that is across cross-section and time series error components. Thus the individual error components are assumed to satisfy the following conditions:

$$E(u_i v_t) = E(u_i w_{it}) = E(v_t w_{it}) = 0$$
$$E(u_i u_j) = 0 \qquad (i \neq j)$$
$$E(v_t v_k) = 0 \qquad (t \neq k)$$
$$E(w_{it} w_{ik}) = E(w_{it} w_{jt}) = E(w_{it} w_{jk}) = 0 \; (i \neq j;\ t \neq k)$$

In accordance with the model, six different regressions were run and the results are reported in columns 1 to 6 of Table A1 in Appendix A. Column 1 reports the results obtained by using all the variables listed in Table 1, while the results in column 2 are obtained by dropping from the regression three explanatory variables (*EQRA, UNSLL,* and *LAMTL*). Thereafter, in each succeeding column of Table 2, one explanatory variable is dropped from the regression, based on poor performance. That is, the results in columns 3-6 are obtained by successively dropping from the regression one explanatory variable that appears trivial from the previous estimation. In this way we tried to determine if successive dropping of statistically insignificant variables improves the previous parameter estimates, thereby indicating the presence of multicollinearity among the different financial ratios.

A close look at the results in Table A1 shows that the estimated coefficients are stable across the different regressions. The adjusted R^2 of about 0.51 for virtually all the six regressions and their respective F-values suggest a fairly good fit for a cross-sectional regression. The Durbin-Watson statistic of 1.87 for each regression indicates that serial correlation does not pose any problem. The results from the different regressions are fairly remarkably consistent and robust.

Of all the eleven explanatory variables used in column 1 of Table A1, only four, viz. loans-to-assets ratio (*LAR*), loans-to-deposits ratio (*LDR*), net income-to-assets ratio (*NITA*), and expenditure-to-liabilities ratio (*EXLR*), are statistically significant and with the expected signs. The estimated coefficients of the remaining seven explanatory variables are not statistically significant, and four of them have the wrong sign. The results are fairly in accord with the results reported by previous studies that also experimented first with a large number of financial ratios (see Whalen and Thomson, 1988; Iyoha and Udegbunam, 1998; Cole and Gunther, 1998). However, the pair-wise correlation coefficients suggest that some of the independent variables are inter-correlated. To alleviate the problem of multicollinearity, and following Farrar and Glauber (1967), the size of the initially over-parameterised model is reduced by dropping three of the seven variables that performed poorly.

The results in column 2 of Table A1 show that the dropping of these variables has not significantly improved the results. The same four variables found significant in column 1 are statistically significant in column 2, but with slightly improved adjusted-R^2 and F-ratio. As already noted, the four variables that are statistically insignificant in column 2 are sequentially dropped (i.e. one after the other) until the preferred (more parsimonious) model, whose estimation results are reported in column 6 and in Table 2, is

achieved. The model in Table 2 is more parsimonious because it achieved maximum goodness-of-fit with minimum number (only four) of the explanatory variables. It is striking to observe that the four explanatory variables (*LAR, LDR, NITA* and *EXLR*) are consistently significant with the expected signs in all the six regressions in Table A1.

Table 2. Estimation results from pooled 1995-1997 data

Variable	Coefficient	T–Statistic
Intercept	–0.1143	–3.453***
LAR	0.4107	2.864***
LDR	0.1887	2.300**
NITA	–0.6840	–2.515**
EXLR	0.6566	3.492***
Adj-R^2 = 0.51	F-Statistic = 13.42	
S.E = 0.97	D.W = 1.88	

Notes: ** and *** indicate significant at the 5% level and at the 1% level respectively.

The regression results consistently indicate that differences in management quality and level of credit risk account for most of the diversity in loan problems and losses among commercial banks in Nigeria, in the mid-1990s. That is, loan loss variations among commercial banks are largely a reflection of differences in management quality and management's risk preferences. Of the three financial ratios used to proxy management quality, the empirical evidence shows that two (*NITA* and *EXLR*) have individually a strong influence on the banks' loan portfolio performance, and therefore on the size of the banks' problem loans and loan losses.

The consistently negative and significant relationship between earnings (*NITA*), or return on assets (*ROA*), and commercial banks' problem loans and losses, implies that an increase in earnings would most likely result in a significant reduction in problem loans and loan losses. On the other hand, a persistent decline in return on assets, which usually characterises an inefficiently managed bank, would tend to motivate the managers to operate a riskier bank in order to increase earnings. This evidence is consistent with our proposition that falling profitability due to poor earnings increases the likelihood of taking excessive credit risks by bank managers, with the consequences of large loan losses.

Again, in accordance with our proposition, the results consistently indicate a positive relationship between operating costs (*EXLR*) and loan losses. With high operating costs (cost inefficiency) and consequent fall in earnings,

inefficient bank managers tend to exhibit an excessively tolerant attitude towards risk. This was why as operating costs sky-rocketed and earnings sharply fell with deregulation and the phenomenal increase in the number of banks, a large number of poorly managed banks failed. Most of these banks failed mainly because of their increased motivation to take excessive risk and the attendant surge in loan losses (see Doguwa, 1996; Udegbunam, 1999).

Perhaps not surprisingly, the coefficient on the ratio of profit-to-earning assets is not statistically significant, suggesting that this ratio is not a good indicator of management efficiency. It is most likely that risk-averse managers trade off low profits for reduced credit risk. That is, these banks tend to accept lower profits for minimum exposure to credit risk. To reduce the probability of default, these banks try to monitor their loan customers with greater diligence and thus incur high operating costs (see Jordan, 1998b).

Interestingly, the two financial ratios (*LAR* and *LDR*) used to measure credit risk passed the t-test at the 5% level in all the regressions, with the expected positive signs. Clearly, banks that take excessive credit risks are much more likely to encounter serious loan problems than banks that are moderately aggressive in their loan portfolio. The higher are these two financial ratios the more likely is high credit risk and the greater are the chances of increased problem loans and loan losses. Recent studies by Iyoha and Udegbunam (1998) and Udegbunam (2000) show that asset quality problems, and an inherent danger of excessive risk taking, were the key factors in bank failures in Nigeria in the early 1990s. During this period lending standards fell with heightened competition occasioned by deregulation. In order to bolster their sharply declining earnings, banks, especially the new ones, tended to exhibit an excessively tolerant attitude towards risk.[2] However, the evidence provided in Appendix B is largely mixed or inconclusive, as to whether new banks are more risk loving than old banks. Nevertheless, the findings of the above studies suggest that a large number of especially new banks were financially distressed and many failed, as loan losses and charge-offs escalated.

It is somewhat surprising that none of the four financial ratios (*UNLTA,*

---

[2] With the tremendous increase in the number of new banks since deregulation, and consequent stiff competition and sharply falling profitability, it is expected that these new banks would be more inclined to take excessive risk in an attempt to bolster earnings.

Using a simple logit regression model, an attempt is made in Appendix B to verify if new banks are more prone to credit risk than old banks. The evidence on risk preferences of these two sets of banks appears to depend on how the credit risk is measured (see Appendix B for details).

LLPEQ, UNSLL and *LAMTL*) used to proxy credit policy is statistically significant. The successive dropping of these variables (ratios) from the regression models has not significantly improved the performance of any of these variables. This rules out multicollinearity as the source of the poor performance of these measures of credit policy. The inability of any of these ratios to explain the diversity in loan problems and losses among the commercial banks may be because:

- The financial ratios used are not appropriate proxies for credit policy. Since collateralising a loan does not guarantee repayment, unsecured loan ratios used may not be appropriate proxies for lending policy. Also on account of the banks' other ulterior motives for loan-loss provisions (see Docking, Hirschey and Jones, 1997; Gunther and Moore, 2000), loan-loss provision ratio may not be a good proxy for lending policy. The problem is that alternative financial ratios are not readily available and that explains the use of four different ratios, in an attempt to achieve a better result.
- There are no significant differences in lending policies; these banks operate in the same financial, regulatory and policy environment, and have barely finished transiting from a highly regulated to a deregulated financial environment. Under the regulated environment they operated with credit guidelines and credit ceilings.
- The differences in credit policy, if they exist, may have been fully reflected in other ratios measuring bank performance, including measures of credit risk. Certainly a good credit policy, if strictly followed, will manifest in low credit risk, while poor credit policy will result in high credit risk, liquidity risk, capital and ultimately insolvency risk. This hypothesis is tested and results are reported in Appendix C. The results do not appear to support the hypothesis, rather it seems that the banks still operate under similar credit policies (see Appendix C for a fuller discussion).

The evidence provided by these results suggests that capital adequacy, measured by equity/risk assets (*EQRA*), and leverage, measured by borrowings/equity (*BER*), do not offer any significant help to the explanation of the variations in loan problems and losses among commercial banks in Nigeria. It is clear from the results that none of these two ratios exerts any significant influence on the size of problem loans and loan losses at these banks.

The finding of this study is not completely in conformity with the published results of similar studies in other countries. In particular, the

studies by Huh and Kim (1994) for Korea and Japan, and Jordan (1998a) for United States (U.S.), found evidence that is markedly different from the one reported in the present study. Undoubtedly, the banking environment of any country has its unique characteristics. For example, there are differences in accounting methods, organizational structure, legal and regulatory framework, and training and personnel. These differences are reflected in the financial ratios used in bank performance studies. Thus, a cross-country comparison of bank performance studies should not be unmindful of these differences, as these financial ratios may not be strictly comparable across independent countries.

It may not be surprising therefore that Jordan finds evidence which suggests that deficient management skills or poor management quality could not offer any significant help to the explanation of the problems U.S. banks in New England have with their loan portfolios. His results show that no significant relationship exists between cost efficiency and problem loans.

Also contrary to the present study, Jordan's results indicate that lending policies are important determinants of the amount of problem loans in U.S. banks. He finds a highly significant relationship between the measures of lending policies and the problem loans at the U.S. banks. Not surprisingly, his evidence on the effect of credit risk is in perfect conformity with the results presented here. He finds that credit risk is a major determinant of loan losses in U.S. banks.

Thus, except for credit risk, which is in accord with our findings, Jordan's results are in sharp contrast with the findings of this study. While management efficiency and credit risk are found to be the major factors determining loan problems in Nigeria, Jordan finds lending policies and credit risk as the major forces driving loan problems in U.S.A.

The methodology adopted in the study by Huh and Kim is significantly different from that used in the present study. It is an aggregative time-series study that focussed more on exogenous macroeconomic factors than on internal financial factors. Their study included only three financial factors: leverage ratio, bank borrowing/total liabilities ratio, and loan growth rate. Of the three financial factors, only two: viz., leverage and borrowing/total liabilities, measuring credit risk, are found to offer significant explanations to changes in the bad loan rate in Japan. In the case of Korea, only leverage is found to be statistically significant. Four out of the five macroeconomic variables are found to exert strong influence on the bad loan rate in Japan, but in Korea they exert only a marginal influence. The findings of the study by Huh and Kim are remarkably different from the evidence provided in the present study. The difference should be attributed largely to the methodology,

financial ratios used, and the economic and financial environment of the countries studied.

5. *Policy Implications*

The empirical evidence appears to suggest that banks responded to declining profitability and shortage of qualified bankers by pursuing riskier strategies in their lending, staff recruitment and other personnel policies. Some of the banks, in particular the new ones, aggressively competed for qualified bankers by offering salaries far above the industry average, and this mounted further upward pressure on operating cost. The increased deterioration of the loan quality, mounting loan losses, and the fragility of the entire banking system, stemming from high credit risk and poor management quality, have wide implications for management, and official regulatory and supervisory policies.

A first step to achieving a safe and sound banking practices that will minimize loan losses, increase efficiency, and promote financial stability, is to install strong corporate governance. A bank's board of directors has many important responsibilities. These responsibilities include hiring of the bank's management team, formulating policies, setting objectives, and ensuring management compliance with policies. Thus the installation of a strong board of directors is undoubtedly a first line of defence against potential banking problems, including credit quality problems.

In addition, there is need for a constant evaluation and strengthening of the supervisory framework for assessing management competence. In this ever changing banking and financial environment, management-focussed supervisory framework should be flexible and adaptable. With adequate supervisory framework and regular monitoring, banks will be encouraged to have on their boards people with good education, relevant experience, integrity and courage. The members of the board should not only be experienced and active, they should also have vested interest by means of a strong ownership position in the banks. With greater financial stakes, members of the board would most likely have greater motivation to complement official supervision by closely monitoring and supervising the banks' management.

The board and senior management should set policies that will ensure that potential employees of the bank are adequately screened for a good and relevant education, a track record, and integrity. A good personnel policy could encourage efficiency and integrity; it should make provision for

training, adequate compensation, and other performance incentives. The bank senior management should ensure that adequate compensation, performance incentives, and the entire corporate culture go down the lines of management.

The strong evidence of high credit-risk provided by this study appears to suggest that both internal credit risk management and official supervisory framework for assessing credit risk exposure of these banks are not adequate. With the rapidly changing banking market structure and increasing complexity of the loan market, there is need for a more dynamic approach to credit risk management and indeed the entire internal control system. This approach should provide, among other things, a comprehensive guide for standard practices regarding credit risk management, such as credit limits, credit standards or criteria, credit risk monitoring, and audit and control procedures. The regulators and the banks' management should ensure that banks comply with prudential regulations, applicable laws and regulations, and official lending policies and procedures. It is the joint responsibility of the banks' management and regulators to ensure that future credit quality problems are avoided through a sound credit risk management and internal control system.

6. *Conclusion*

The evidence provided by the regression results presented above suggests that differences in management quality and the level of credit risk are the major determinants of the diversity in the severity of the recent surge in problem loans and losses among commercial banks in Nigeria. Of the eleven financial ratios used to measure credit policy, management quality, and three other areas of bank performance (capital adequacy, leverage and credit risk), four are found to be the most critical factors determining the variations in problem loans and losses across the commercial banks. Of the four critical factors, two (*NITA* and *EXLR*), are indicators of management quality and two (*LAR* and *LDR*), are indicators of credit risk. There is therefore, an overwhelming evidence that differences in management quality and credit risk are the major force behind the wide diversity in the severity of loan problems and losses among commercial banks, in the mid-1990s. Thus, the dramatic rise in loan losses at some of these banks and recent wave of financial distress and bank failures are largely the result of poor management quality and increased incentive of the inexperienced bank managers for greater risk taking. On the other hand, the high portfolio and profit performance of an equally large proportion of these banks can be attributed to

high management quality, and minimal exposure of these banks to credit risk and losses. That is, the lending expertise of the managers of these banks, which includes their superior ability to screen loan applicants for credit worthiness, to workout problem loans, and to value and perfect collaterals, translates into lower credit risk, lower workout costs, lower operating expenses, and high profitability.

A somewhat surprising finding is that differences in credit policy do not seem to make any significant contribution to the variations in loan problems and credit losses among commercial banks in Nigeria. None of the four financial ratios used to proxy credit policy is found to exert any significant influence on the loan-loss rate of these banks. This apparently implausible finding appears to suggest that;

- credit policy is not properly measured by any of the four financial ratios used;
- if these proxies are adequate, then it may be that there are no significant differences in the lending policies of these banks; and
- if there are differences in lending policies, they may have been fully reflected in the other measures of bank performance, including the measures of credit risk.

Another notable finding is that neither capital adequacy nor financial leverage could make any significant contribution to the explanation of the variations in loan problems and losses across the commercial banks, in the mid-1990s. None of the two financial ratios measuring capital adequacy and leverage exerts any significant effect on the banks' loan losses. While this evidence appears to conform with the argument by several studies (see for example Jones and King, 1992, 1995; Peek and Rosengren, 1996), that capital ratios are lagging indicators of banks problems, it contradicts the suggestion that managers of poorly capitalized banks tend to be more aggressive lenders with high default risk (see Kane, 1989, Keeton, 1999).

APPENDIX A

Table A1. Estimation results

	1	2	3	4	5	6
Intercept	−.1409** (−2.345)	−.1504** (−2.853)	−.1500*** (−3.667)	−.1204*** (−3.516)	−.1148*** (−3.524)	−.1143*** (−3.453)
LAR	.3622** (2.127)	.3858** (2.421)	.3852** (2.627)	.3897** (2.640)	−3747** (2.603)	.4107** (2.864)
LDR	.2122** (2.405)	.2087** (2.465)	.2087 (2.485)	−2204** (2.624)	.2183** (2.607)	.1887** (2.30)
NITA	−.5655* (−1.79)	−.5895** (−2.004)	−.5887** (−2.097)	−.6478** (−2.324)	−.6263** (−2.282)	−.6840** (−2.515)
EXLR	.7690** (2.888)	.8123*** (3.547)	.8117*** (3.705)	.6796*** (3.486)	.6506*** (3.494)	.6566*** (3.492)
UNLTA	−.0844 (−1.033)	−.0846 (−1.056)	−.0849 (−1.109)	−.0907 (−1.180)	−.0908 (−1.186)	
LLPEQ	−.0035 (−1.007)	−.0036 (−1.085)	−.0036 (−1.323)	−.0011 (−.556)		
BER	.0036 (1.022)	.0038 (1.111)	.0038 (1.304)			
PTEA	.0079 (.078)	.0010 (.011)				
EQRA	−.0209 (−.389)					
UNLL	.0282 (.468)					
LAMTL	−.0134 (−.352)					
Adjusted R^2	0.48	0.51	0.52	0.51	0.52	0.51
S.E	0.10	0.98	0.97	0.98	0.97	0.97
F–Statistic	6.13	8.31	9.38	10.23	11.77	13.42
D.W	1.87	1.87	1.87	1.86	1.87	1.88

Notes: *, ** and *** indicate significant at the 10% level, at the 5% level, and at the 1% level respectively.

APPENDIX B

Evidence on Risk-Taking of New Banks

The simplest way to test if new banks are more risk loving or have greater incentive for risk taking, is to use a binary choice model (BCM). The binary variable takes discrete outcomes: new bank or old bank. With this model we estimate the likelihood that the bank with the greatest motivation to assume risk is a new bank.

Specifically, we use a logit model to evaluate how credit risk ratios can predict that the bank with the greatest incentive to take risk is a new bank. Thus, the dependent variable is the probability of the bank being a new bank, and the explanatory variables are leverage ratio (*BER*), capital adequacy ratio (*EQRA*), credit risk ratios (*LAR* and *LDR*), the ratio indicating management quality (*NITA*), and a measure of credit policy (*UNLTA*). The focus is on the credit risk ratios; the other financial ratios are included as control variables.

The logit model estimated is of the following form

$$Z_i = \log \frac{P_i}{1 - P_i} = \alpha + \beta x + e_i$$

where Z_i is the binary variable that has the value one if bank i is a new bank and zero otherwise. It is the logarithm of the odds that bank i is a risk lover; β is a vector of parameters to be estimated; x is a vector of independent variables; α is a constant, and e_i is the error term. The model is estimated using the pooled data covering the period 1995-1996, and with the original 24 sample banks, twelve of which are new banks.

Table A2. Logit regression: estimation results from pooled 1995-1996 data. Dependent variable: probability of being a new bank

Variable	Coefficient	T-Statistic
Intercept	0.9569	0.851
BER	0.0615	1.124
EQRA	−1.5806	−1.199
NITA	9.0424	1.114
LDR	8.8596	2.356**
LAR	−19.5522	−2.704**
UNLTA	−2.3114	−0.976
Goodness of fit measures:		
Pseudo-R^2	0.1799	
Percentage of correct prediction	71	
Log-likelihood function	−27.822	
Number of correct predictions	17	
Degrees of freedom	42	

Notes: ** indicate statistically significant at the 5% level.

The estimation results presented in Table A2 indicate that the key predictive variables are the two financial ratios measuring credit risk (*LAR* and *LDR*). Each of the coefficients on these two measures of credit risk passes the t-test at the 5% level, but surprisingly the coefficient on *LAR* (Loans/Assets ratio) has a negative sign which is counter intuitive. The goodness-of-fit measures show a fairly impressive overall correct prediction of 71%, indicating that of the 24 sample banks, 17 are

correctly classified.

Interestingly, the coefficient on *LDR* (loans/deposits ratios) is statistically significant with the expected positive sign. The result shows that the higher the loans/deposits ratio (the higher the credit risk), the more likely the bank is a new bank. That is, banks that take excessive credit risk by operating with high loans/deposits ratio have higher probability of being new banks.

The implication of the significant negative coefficient on *LAR* is that the higher the ratio of loans/assets, the less likely the banks is a new bank, or the more likely the bank is an old bank. This suggests that, although the two ratios (*LAR* and *LDR*) measure credit risk, *LDR* is a more appropriate measure if the bank is a new bank, and *LAR* a better measure if the bank is an old bank. The overall evidence is therefore somehow mixed; it appears to depend on how the credit risk is measured. When credit risk is measured as loans/deposits ratio, new banks are found to exhibit greater love for risk taking than old banks. When it is measured as loans/assets ratio, older banks are found to exhibit greater tolerance attitude towards risk than new banks.

APPENDIX C

Evidence on the Hypothesis that Credit Risk Reflects Differences in Credit Policy

To verify the suggestion that the differences in credit policy may have been fully reflected in other financial ratios, especially the measures of credit risk, we regressed the two measures of credit risk (*LDR* and *LAR*) individually on the four indicators of credit policy (*UNLTA, UNLL, LLPEQ* and *LATML*). The results obtained (see Table A3, column 1) from this estimation indicate that there is little or no relationship between the four credit policy variables taken together, and any of the two indicators of credit risk. This evidence is clear from the adjusted-R-Squared which shows that only 2% of the differences in *LDR*, and 14% in *LAR* could be explained by the respective models. The F-statistics of 1.071 for *LDR* model and 1.573 for *LAR* model are not significant at any acceptable level of confidence. Although the variable *UNLTA* is statistically significant in the two models, the extremely poor fit suggests that this result should not be relied upon for any reasonable inference.

In order to obtain a better result, the model is extended by including the measure of leverage (*BER*), and the two measures of management quality (*NITA* and *PTEA*), as control variables. Although the results obtained with the extended model (see Table A3, column 2) are clearly more robust, as indicated by adjusted-R^2, F-statistics and some of the t-ratios, the coefficients on all the four indicators of credit policy are again consistently insignificant. This further confirms that none of the four indicators of credit policy can explain any variation in credit risk. The results show that with falling profitability, banks tend to assume greater credit risk.

The implication of our finding is that there are no significant differences in the credit policy of the sample banks; it seems they operate as if they are still under the regulatory credit guidelines and ceilings. But it may also be that the credit policies

are not forcefully binding on the managers of these banks; it seems the managers have a large room to use their own initiatives in credit decisions. This inference is supported by the highly significant coefficient on *PTEA* measuring management quality which implies that management influence is dominant.

Table A3. Regression results

Dependent variable: *LDR* (proxy for credit risk)			Dependent variable: *LAR* (proxy for credit risk)		
	Coefficients			Coefficients	
Variables	(1)	(2)	Variables	(1)	(2)
Intercept	0.2607***	0.3358**	Intercept	0.1734**	0.2196***
	(3.096)	(2.828)		(2.628)	(3.094)
UNLTA	2.1966**	0.7949	*UNLTA*	1.6714**	0.4497
	(2.249)	(0.825)		(2.556)	(0.780)
UNLL	–0.2256	0.2160	*UNLL*	–0.3179	–0.0165
	(–0.633)	(0.671)		(–1.442)	(–0.086)
LLPEQ	–0.0039	0.0090	*LLPEQ*	–0.0029	0.0037
	(–0.411)	(0.966)		(–0.0490)	(0.667)
LATML	0.05704	–0.1348	*LATML*	0.0657	–0.0612
	(0.442)	(–1.189)		(0.836)	(–0.901)
BER		–0.0109	*BER*		–0.0084
		(–1.177)			(–1.504)
NITA		0.9369	*NITA*		0.8626**
		(1.437)			(2.212)
PTEA		0.8147***	*PTEA*		–0.5030***
		(–3.355)			(–3.465)
Adjusted-R^2	0.019	0.434	Adjusted-R^2	0.141	0.544
F-Statistic	1.071	3.516	F-Statistic	1.573	4.916
D.W	1.860	1.857	D.W	1.826	1.915

Notes: ** and *** indicate statistically significant at the 5% level and at the 1% level respectively.

REFERENCES

BERGER A.N. and DEYOUNG R., "Problem Loans and Cost Efficiency in Commercial Banks", *Journal of Banking and Finance*, 1997, *21*, pp. 849-70.

COLE R.A. and GUNTHER J.W., "Predicting Bank Failures: A Comparison of On- and Off-Site Monitoring Systems", *Journal of Financial Services Research*, April 1998, *13*, pp. 103-17.

DEMIRGUC K.A. and DETRAGIACHE E., "The Determinants of Banking Crises: Evidence from Industrial and Developing Countries", Policy Research Working Paper No. 1828, Washington, D.C.: World Bank, 1997.

DOCKING D.S., HIRSCHEY M., and JONES E., "Information and Contagion Effects of Bank Loan-Loss Research Announcements", *Journal of Financial Economics*, February 1997, *43*, pp. 219-39.

DOGUWA S.I., " On Early Warning Models for the Identification of Problem Banks in Nigeria", *CBN Economic and Financial Review*, March 1996, *34*(1), pp. 462-87.

EDWARDS F.R. and MISHKIN F.S., "The Decline of Traditional Banking: Implications for Financial Stability and Regulatory Policy", *Federal Reserve Bank of New York Economic Policy Review*, 1995, *1*(2), pp. 27-45.

EMMONS W.R, "Increased Risk-taking Versus Local Economic Conditions as Causes of Bank Failures", Proceedings of the 29th Annual Conference on Bank Structure and Competition, Federal Research Bank of Chicago, 1993, pp. 189-209.

ESTRELLA A., "A Prolegomenon to Future Capital Requirements", *Federal Reserve Bank of New York Economic Policy Review*, July 1995, *1*(2), pp. 1-12.

—— et al., "Capital Ratios as Predictors of Bank Failure", *Federal Reserve Bank of New York Economic Policy Review*, July 2000, *6*(2), pp. 33-52.

FARRAR D.E. and GLAUBER R.R., "Multicollinearity in Regression Analysis: The Problem Revisited", *Review of Economics and Statistics*, 1967, *49*, pp. 92-107.

FULLER W.A. and BATTESE G.E., "Estimation of Linear Models with Crossed-Error Structure", *Journal of Econometrics*, May 1974, *2*, pp. 67-78.

FURLONG F.T., "Capital Regulation and Bank Lending", *Federal Reserve Bank of San Francisco Economic Review*, No. 3, 1992, pp. 23-33.

GILBERT R.A., "Supervision of Undercapitalised Banks: Is There a Case for Change?", *Federal Reserve Bank of St. Louis Review*, May/June 1991, *73*(3), pp. 16-30.

—— et al., "The New Bank Capital Adequacy Standards", *Federal Reserve Bank of St. Louis Review*, May 1985, *67*(5), pp. 12-20.

—— et al., "The Role of Supervisory Screens and Econometric Models in Off-Site Surveillance", *Federal Reserve Bank of St. Louis Review*, Nov/Dec 1999, *81*(6), pp. 31-56.

GUNTHER J.W. and MOORE R.R., "Financial Statements and Reality: Do Troubled Banks Tell All?". *Federal Reserve Bank of Dallas Economic and Financial Review*, Third Quarter 2000, pp. 30-35.

HUH C. and KIM S.B., "Financial Regulation and Banking Sector Performance: A Comparison of Bad Loan Problems in Japan and Korea", *Federal Reserve Bank of San Francisco Economic Review*, No. 2, 1994, pp. 18-29.

HUNTER W.C. and SRINIVASAN A., "Determinants of De Novo Bank Performance", *Federal Reserve Bank of Atlanta Economic Review*, March-April 1990, pp. 14-25.

IYOHA M.A. and UDEGBUNAM R.I., "Prediction of Bank Failure in Nigeria Using Financial Ratios and a Logit Regression Model", PRAJNAN, *Journal of Social and Management Sciences*, October-December 1998, *27*(3), pp. 291-306.

JONES D.S., "The Implementation of Prompt Corrective Action: An Assessment", *Journal of Banking and Finance*, June 1995, *18*(3-4), pp. 491-510.

—— and KING K.K., "The Implementation of Prompt Corrective Action", in *Credit Markets in Transition*, Proceedings of the 28th Annual Conference of Bank Structure and Competition, Federal Reserve Bank of Chicago, May 1992, pp. 68-100.

JORDAN J.S. (1998a), "Problem Loans at New England Banks, 1989 to 1992: Evidence of Aggressive Loan Policies", *New England Economic Review*, January-February 1998, pp. 23-36.

—— (1998b), "Resolving a Banking Crisis: What Worked in New England", *New England Economic Review*, September-October 1998, pp. 49-62.

KANE E.J., "The High Cost of Incompletely Funding the FSLIC Shortage of Explicit Capital", *Journal of Economic Perspectives*, 1989, *3*, pp. 31-47.

KEETON W.R., "Does Faster Loan Growth Lead to Higher Loan Losses?", *Federal Reserve Bank of Kansas City Economic Review*, Second Quarter 1999, *84*(2), pp. 57-75.

—— and MORRIS C.S., "Why Do Bank Loan Losses Differ?", *Federal Reserve Bank of Kansas City Economic Review*, May 1987, *72*, pp. 3-21.

KOCH T.W., *Bank Management*, New York: The Dryden Press-Harcourt Brace Jovanovic College Publishers, 1992.

MITCHELL W.C., "The Decline in the Ratio of Banking Capital to Liabilities", *Quarterly Journal of Economics*, August 1909, *23*, pp. 697-713.

PANTALONE C.C. and PLATT M.B., "Predicting Commercial Bank Failure Since Deregulation", *New England Economic Review*, July/August 1987, pp. 37-47.

PEEK J. and ROSENGREN E.S., "The Use of Capital Ratios to Trigger Intervention in Problem Banks: Too Little, Too Late", *New England Economic Review*, September-October 1996, pp. 49-58.

RANDALL R.E., "Lessons from New England Bank Failures", *Economic Review*, May-June 1993, pp. 13-38.

SACHS J. et al., "Financial Crises in Emerging Markets: The Lessons from 1995", *Brookings Papers on Economic Activity*, No. 1, 1996, pp. 147-98.

UDEGBUNAM R.I., "Examining the Causes of Bank Failure in Nigeria Since Deregulation: Some Empirical Results and Implications for the Policy", *Vision, The Journal of Business Perspectives*, July-December 1999, *3*(3), pp. 8-17.

——, "Financial Distress and Performance Differences Among Commercial Banks in Nigeria: A Multivariate Ratio Analysis", *Finance India, The Quarterly Journal of Finance*, June 2000, *15*(2), pp. 551-66.

WHALEN, G. and THOMPSON J., "Using Financial Data to Identify Changes in Bank Condition", *Federal Reserve Bank of Cleveland Economic Review*, Quarter 2, 1988, pp. 17-26.

WHEELOCK D. and WILSON P.H., "Explaining Bank Failures: Deposit Insurance, Regulation, and Efficiency". *Review of Economics and Statistics*, 1995, pp. 689-700.

WHITE H., " A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity", *Econometrics*, 1980, *48*, pp. 817-38.

ABSTRACT

This paper examines the role of management quality and credit policy in the wide disparity in the severity of loan problems and loan losses among Commercial Banks in Nigeria, in the 1990s. Using a simple model of bank loan performance and a pooled time-series cross-section data, six different regressions are run in a stepwise process. The evidence provided by these regressions suggests that management quality and credit risk are the key determinants of the variations in loan problems and losses across the commercial banks in Nigeria. Surprisingly, differences in credit policy have little or no effects on loan-loss diversity in the Nigerian Commercial Banks in the 1990s.

JEL classification: *G2, G21*
Keywords: *management efficiency, loan-loss, commercial banks*

RISEC, Volume 50 (2003), No. 4, 569-580

FOREIGN DIRECT INVESTMENT AND ECONOMIC GROWTH IN SUB-SAHARAN AFRICA

by
ANTHONY E. AKINLO*

1. *Introduction*

The economic development literature is replete with studies on the impact of foreign direct investment (FDI) on economic growth[1]. A generally accepted conclusion is that foreign direct investment played a significant role in promoting economic growth in host countries. Several authors identified ways in which FDI affects growth. These include increased capital formation, improved efficiency of locally owned firms, better technological change and human capital augmentation. However, probably due to the limited capacity of the African countries to attract foreign direct investment for a very long time in the 70s and 80s, not many studies have analysed the impact of FDI in Africa as compared to that in other regions as Asia, and Latin America.[2] As many African countries opened up their economies in the late 80s which resulted in increasing flows of FDI to the continent, it is imperative to assess the impact of these inward flows on its growth.[3] This constitutes the objective

* Department of Economics, Obafemi Awolowo University, Ile-Ife (Nigeria) and Institute of Developing Economies, JETRO (Japan). E-mail: aakinlo@oauife.edu.ng

The financial support and facilities provided by the Institute of Developing Economies, JETRO, Japan, for carrying out this study is gratefully acknowledged. The author is however exclusively responsible for any error or omission that the paper may contain.

[1] For a detailed literature survey, see de Mello (1997) and Akinlo (2003).

[2] Most of the existing studies have focused primarily on the determinants of foreign direct investment in Africa.

[3] It must be pointed out that FDI flows to Africa are still low in magnitude when compared with those of other regions. However, when the present inflows to Africa are related to economic size, the difference is not so significant.

of this paper.

Specifically, the paper examines the effect of FDI on economic growth through an augmented Cobb-Douglas production function in which the externality of FDI is explicitly incorporated. In addition, the paper identifies the channels through which FDI affects growth by specifying separate equations for capital accumulation and total factor productivity.

The paper is organized as follows: Section 2 provides the model specification. Section 3 gives the empirical results and Section 4 contains the conclusions.

2. *The Models*

2.1. *FDI-growth Equation.* - We specify an augmented Cobb-Douglas production function in which the externality (positive or negative effect) of FDI is incorporated.

Taking a cue from previous studies (de Mello, 1997, and Ramirez, 2000, among others), the Cobb-Douglas production function is stated as[4]:

$$G = A(L, K_P, E) = AL^{\alpha}K^{\beta}E^{1-\alpha-\beta} \quad (1)$$

where G is real output, K_P is private capital stock, L is labour and E represents externality generated by additions to the stock or flow of FDI. α and β are the shares of domestic labour and private capital respectively, and A captures the efficiency of production. It is equally assumed that α and β are less than one, but that there are diminishing returns to labour and capital inputs.

The externality, E, can be represented by a Cobb-Douglas function of the type:

$$E = (L, K_P, K_f^{\nu})^{\phi} \quad (2)$$

where ν and ϕ are, respectively, the marginal and the inter-temporal elasticities of substitution between private and foreign capital. We assume $\nu > 0$ so that a larger FDI generates a positive externality to the economy. If $\phi > 0$ intertemporality prevails, and if $\phi < 0$, additions to the stock or flows of FDI crowd out private capital over time and diminish the growth potentials of the host country (Borsworth and Collins, 1999; Ramirez, 2000).

[4] Time and countries subscripts were dropped for convenience.

Combining equations (1) and (2) we obtain:

$$G=AL^{\alpha+\phi(1-\alpha-\beta)}K_P^{\beta+\phi(1-\alpha-\beta)}K_f^{\nu\phi(1-\alpha-\beta)} \tag{3}$$

Taking the log and time derivative of equation (3) we obtain:

$$g_{G_i} = g_A + \{\alpha+\phi(1-\alpha-\beta)\}g_L + \{\beta+\phi(1-\alpha-\beta)\}g_{K_P} + \{\nu\phi(1-\alpha-\beta)\}g_{K_f} \tag{4}$$

where g_i is the growth rate of $i = G, A, L, K_P$, and K_f. Thus, equation (4) states that (given that ν and $\phi > 0$) additions to the flow or stock of FDI will augment the elasticities of output with respect to labour and capital by a factor $\phi(1-\alpha-\beta)$.

From equation (4), the equation estimated is:

$$\Delta\ln G = d_0 + d_1\Delta\ln L + d_2\Delta\ln K_P + d_3\Delta\ln K_f + d_4\Delta\ln X + d_5\Delta\ln C_g + d_6\ln T \tag{5}$$

where the variables are defined thus: C_g is real government consumption expenditures and T is a time trend designed to capture any secular trend in output growth during the period under consideration. L, K_f and K_p are as earlier defined and X is real exports. d_1 and d_2 are expected to be positive, while d_3 can be positive or negative depending on whether foreign capital complements or substitutes private capital formation; d_4 is anticipated to be positive as greater export is associated with greater economic growth through various channels (Tyler, 1981; Ram, 1987; and Feder, 1983). d_5 is indeterminate as it can be positive or negative depending on whether or not government consumption "crowds out" or "crowds in" private consumption expenditures. d_6 can be positive or negative depending on whether or not annual average growth rate in the countries accelerated or decelerated over the period under consideration. Few additional variables are added to determine their effects on the growth level in the selected countries. These are: real government consumption (C_g) and export (X).

Two main channels through which FDI stock (F_s) or flows (F_l) may influence growth are also examined. These are capital formation and total factor productivity.

Thus, we specify and estimate the equation for each with FDI as the independent variable to determine whether FDI influences growth through capital formation or factor productivity, or both.

2.2 *Capital formation equation.* - In specifying the capital formation equation, we assume that aggregate investment is made up of domestic (I_d) and foreign real investment (I_f) undertaken by multinationals. Thus,

$$I_{gd} \equiv I_d + I_f \tag{6}$$

Foreign investment by multinationals is assumed to be a function of foreign direct investment (F).[5] Thus,

$$I_f = h(K_f) \tag{7}$$

Domestic investment is modelled following the neoclassical investment model as exemplified in the work of Hall and Jorgerson (1967). We assume that investment is essentially a stock adjustment variable responding to the difference between the desired and actual capital stock. Investment adjusts partially to this difference because firms face liquidity constraints to investment and because the adjustment takes time. Therefore, the basic equation is stated as follows:

$$I_d = \gamma (K^*_d - K_d) \tag{8}$$

where K^*_d represents the capital stock desired by domestic firm, and $\gamma < 1$. In our model, we assume that the desired level of capital stock depends positively on expected growth G^e, the difference between actual output and full capacity output (y) and the rate of interest. We include interest rate to see its impact on investment particularly with the liberalization of the interest rate in most countries of the continent. Hence, the domestic investment equation becomes:

$$K^*_d = \pi_0 + \pi_1 G^e + \pi_2 y + \pi_3 r \tag{9}$$

where $\pi_1, \pi_2 > 0; \pi_3 < 0$.

Consider the law of motion of the capital stock:

$$K_d = (1 - \delta) K_{d-1} + I_{d-1} \tag{10}$$

[5] Here. we assume that there is no lag between FDI and foreign investment. However, in some cases. the resources that move from one country to another as foreign direct investment are often not used at once to finance real investment. But in view of the fact that we used the stock value of FDI. our assumption may be more reasonable.

where δ is the annual depreciation rate.

If we combine equations (8) through (10), we obtain:

$$(11) \qquad I_d = \varphi_0 + \varphi_1 G^e + \varphi_2 y + \varphi_3 r + \gamma I_{d-1} + \gamma^* I_{d-2}$$

where $\varphi_0 = \pi_0 + \gamma^2(1-\delta)^2 K_{d-2}$, $\varphi_1 = \gamma\pi_1$, $\varphi_2 = \gamma\pi_2$, $\varphi_3 = \gamma\pi_3$, $\gamma^* = \gamma^2(1-\delta)$.

Substituting equations (11) and (7) into equation (6) and collecting the like terms, we obtain our model for domestic total investment given as:

$$(12) \qquad I_{gd} = \varphi_0 + \varphi_1 G^e + \varphi_2 y + \varphi_3 r + hk_f + \gamma I_{d-1} + \gamma^* I_{d-2}$$

We assume adaptive expectation, in which $G^e = G_{-1}$. Hence, we have:

$$(13) \qquad I_{gd} = \varphi_0 + \varphi_1 G_{-1} + \varphi_2 y + \varphi_3 i + hk_f + \gamma I_{d-1} + \gamma^* I_{d-2}$$

Our estimated investment equation is derived from equation (13) and is stated as:

$$(14) \qquad \Delta \ln I_{gd} = \Omega_0 + \Omega_1 \Delta \ln K_f + \Omega_2 \Delta \ln P_c + \Omega_3 \ln i + \Omega_4 \Delta \ln B_g + \Omega_5 \Delta \ln C_g + \Omega_6 \Delta \ln I_{gd-1}$$

where P_c is the growth rate of per capita income, I_{gd} is domestic investment measured as growth rate of gross fixed capital formation, B_g is government budget balance over GDP, C_g is real government consumption expenditures and I_{gd-1} is the lagged dependent variable. Ω_1 can be positive or negative. Ω_2 is expected to be positive while Ω_3 is expected to be negative; Ω_4 can be positive or negative, Ω_5 can be positive or negative while Ω_6 is anticipated to have a positive sign.

2.3. *Total factor productivity and FDI equation.* - We specify and estimate the equation for total factor productivity to see the impact of FDI. Such equation is:

$$(15) \qquad \ln F_p = \vartheta_0 + \vartheta_1 \Delta \ln K_f + \vartheta_2 \ln U_{-1} + e$$

The productivity growth rate (F_p) is calculated for sample countries as Solow residuals defined in the Appendix. The variables included are stock of FDI and unemployment rate (U), lagged one period as a measure of cyclical

effects in productivity growth. However, we try with a trend variable in view of the discrepancies observed in the unemployment figures for some countries. ϑ_1 can be positive or negative based on our earlier argument and ϑ_2 is expected to be negative.

3. *Empirical Results*

3.1. *Estimation techniques and presentation of the estimation.* - The results of the estimated equations 5, 14 and 15 are presented in Tables 1, 2 and 3 respectively. The equations are estimated using pooled annual data for the period 1990-2001 for twelve countries selected for the study allowing for fixed effects.[6] This allows the intercept term to be estimated separately for each country thereby abstracting from country differences in the long-term growth, even if slope coefficients are constrained to be the same across countries.

FDI and economic growth

The results of the GDP growth rate are as shown in Table 1. The regression results 1-3 use the contemporaneous, first and second lagged stock FDI respectively. The results show that the immediate impact of changes in private capital stock is positive and statistically significant. Using regression 2, a one-percentage point increase in private capital increases the economic growth rate by 0.16%. However, the value reduces to 0.05 with the introduction of lagged values of foreign capital. This tends to show that the two variables are a little bit offsetting as the significance of domestic private capital vanishes with the introduction of foreign capital. The contemporaneous changes in the employment growth rate have a positive sign but not significant. The FDI variable (FDI stock) has a positive sign when lagged twice. This result should not come as a surprise because the positive externalities generated in the form of a greater transfer of technology and managerial know-how among others, are likely to impact on growth in the long term. The result in formulation 3 shows that a one-percentage point increase in foreign capital stock will increase growth by 0.18%. The coefficient of government consumption shows that consumption tends to crowd out private investment with adverse effect on output. However,

[6] The twelve countries included in our studies are Botswana, Côte d'Ivoire, Egypt, Ghana, Kenya, Mauritius, Morocco, Nigeria, South Africa, Swaziland, Tunisia, Zimbabwe. For an unbiased choice, we include only the countries with stock markets and whose stock markets were established in or before 1990.

conclusive evidence cannot be inferred from it since the coefficient is not significant. The growth rate of export has a positive and significant effect. Using formulation 3, the result shows that a one-percentage point increase in export increases growth by 0.29%. The result shows that the coefficient of export has increased with the reforms when compared with the 0.11 – 0.13 values obtained in most previous studies (Fosu, 2000). The coefficient of capital formation has the expected positive sign and is significant. The trend variable shows up with a positive sign but is not significant. The positive coefficient of the trend variable might partly reflect the increase in growth experienced by most African countries after the reforms.

Table 1. Fixed effects panel regression. Dependent variable: $\Delta \ln G$ ($N = 144$)

OLS regression			
Variables	(1)	(2)	(3)
$\Delta \ln F_x$	–0.05 (–0.73)		
$\Delta \ln F_{x-1}$		–0.04 (–0.04)	
$\Delta \ln F_{x-2}$			0.18 (2.23)
$\Delta \ln L$	0.09 (0.06)	0.11 (0.64)	0.20 (0.03)
$\Delta \ln K_p$	0.16 (1.83)	0.16 (1.64)	0.05 (0.78)
$\Delta \ln C_g$	–0.03 (–0.52)	–0.02 (–0.31)	–0.01 (–0.13)
$\Delta \ln X$	0.35 (4.57)	0.35 (4.35)	0.29 (3.35)
$\Delta \ln I_{gd-1}$			0.21 (3.27)
T	0.001 (0.38)	–0.002 (–0.08)	0.01 (1.34)
$\bar{R}^2$	0.22	0.32	0.32

Note: t-statistics are in parentheses.

Capital formation equation

The result of the capital formation equation is as shown in Table 2. Regression 1 uses FDI flows while formulations 2 and 3 both use stock FDI. However, formulation 3 uses the lagged dependent variable as additional explanatory variable. In the three formulations, the two measures of FDI have a positive sign but only the stock variable is significant. Equation 3 shows that a one-percentage point increase in stock FDI increases capital formation by 0.19%. The per capita growth variable has the expected positive sign and is significant. Formulation 3 indicates that a one-percentage point increase in per capita income increases capital formation by 0.7%. Interest rate has a negative sign and is significant. The results in Table 2 unanimously indicate that a one-percentage point increase in interest rate reduces investment by

0.01%. This should not come as a surprise because for many African countries, the interest rate increased phenomenally after the reforms in the mid-80s. This might have increased the cost of borrowing with adverse effects on investment. This result is in line with previous ones that have demonstrated that a high interest rate discourages investment in developing countries (Akinlo, 2001; Odedokun, 1992).

Government budget balance over GDP has a negative sign and is significant. The results show that tight fiscal policy tends to adversely affect investment during the period under consideration. Moreover, the perverse effects of high fiscal deficits have been noted in the literature. They increase the rate of inflation and uncertainty which may reduce investment and economic growth (Fisher, 1993; Barro, 1996; Bruno and Easterly, 1998). The lagged dependent variable has a negative but not significant sign, therefore no definite assertion can be based on it.

Table 2. Fixed effects panel regressions. Dependent variable: $\Delta \ln I_{gdi}$ ($N = 144$)

OLS regression			
Variables	(1)	(2)	(3)
$\Delta \ln F_t$	0.01 (1.16)		
$\Delta \ln F_s$		0.17 (1.76)	0.19 (1.85)
$\Delta \ln P_c$	0.62 (2.71)	0.72 (3.38)	0.72 (3.21)
ΔB_g	−0.01 (−2.52)	−0.01 (−2.88)	−0.01 (−2.19)
I	−0.01 (−2.59)	−0.01 (−3.18)	−0.01 (−2.79)
$\Delta \ln I_{gdi-1}$	−0.02 (−0.28)		−0.04 (0.42)
$\bar{R}^2$	0.24	0.29	0.26

Note: t-statistics are in parentheses.

Total factor productivity

The result of the productivity growth equation is as shown in Table 3. Regressions 1 and 2 use FDI flow and stock values respectively. Both results show that FDI variables have positive signs but significant only with stock FDI. This indicates that FDI increases productivity in the countries included in the study. The result shows that a one-percentage point increase in foreign capital increases productivity by 3.8%. It shows that FDI has spillover effects on African economies. The unemployment variable has a positive sign as against the expected negative sign. However, it is not significant. The lagged values of residuals have a positive sign and are significant. When we replace unemployment with the trend variable as shown in formulation 3, the result

remains the same except that the significance of FDI stock on productivity is reduced. This probably shows that productivity declines over time.

Table 3. Fixed effects panel regression. Dependent variable: $\Delta \ln F_i$ ($N = 144$)

OLS regression			
Variables	(1)	(2)	(3)
$\Delta \ln F_i$	0.03 (1.10)		
$\Delta \ln F_{\backslash}$		3.81 (1.57)	9.71 (1.32)
ΔU_{-1}	0.17 (0.76)	0.17 (0.75)	
$\Delta \ln F_{p-1}$	0.44 (2.04)	0.49 (2.26)	0.52 (2.45)
T			0.28 (1.37)
$\bar{R}^2$	0.45	0.46	0.47

Note: t-statistics are in parentheses.

4. *Conclusion*

The results of our analysis show that private and foreign capital stocks as well as export have positive and significant effects on growth. Moreover, the results show that FDI stocks work through capital formation and factor productivity, though the coefficient of the latter is barely significant. Per capita income enhances capital formation while a tight fiscal policy and high interest rate have negative and significant effects on capital formation. For policy perspectives, the evidence in the paper suggests the need to attract more FDI to Africa through greater reforms. This will involve eliminating all the institutional and legal barriers that tend to discourage foreign investments in the continent. Moreover, the results demonstrate the need to enhance per capita income in the continent while maintaining easy fiscal and monetary policies in the form of a low tax regime and interest rate. This simply suggests that low tax regime and low interest rate coupled with a favorable investment environment will enhance investment inflows, with a positive impact on capital formation and factor productivity, as well as growth in Sub-Saharan Africa.

APPENDIX

Sources of Data

- International Monetary Fund, *International Financial Statistics Yearbook* (various years)
- World Bank, *World Development Indicators* (various years)
- World Bank: *African Development Indicators* (various years)
- African Development Bank, *Selected Statistics on African Countries* (2000)

Description and Measurements of Variables

G = real output (measured as real GDP)
K_p = private capital stock (private investment)
K_f = stock of foreign investment, the estimates of K_p and K_f were generated using a standard perpetual inventory model of the form: $K_t = K_{t-1} + I_t - \delta K_{t-1}$ where K_{t-1} is the stock of capital at time t–1. I_t is the flow of gross investment during period t, and δ is the rate at which private and foreign capital depreciates in period t – 1. In this work, the initial stocks of private and foreign capital were estimated by aggregating over 5 years of gross investment 1986–90 with 5 percent rate of depreciation
F_t = foreign capital flows
L = economically active labour force
X = real export
P_c = per capita income
B_g = government budget balance over GDP
I_{gd} = domestic investment measured as growth rate of gross fixed capital formation
C_g = real government consumption
i = interest measured as the discount rate
F_p = total factor productivity measured as Solow's residual defined as: $F_p = G - 0.4K - 0.6L$ where G is real output, K is growth rate of real gross capital (using real fixed capital formation) and L is the labour force. We assume constant labour share of 0.6 considering the abundance of labour as against scarce capital in these economies.
U = is the unemployment rate
T = time trend

REFERENCES

AKINLO A.E., "An Alternative Framework of Assessing the Impact of Adjustment Policies on Industrial Performance: An Application to Time Series Data for Nigeria", *RISEC (Rivista Internazionale di Scienze Economiche e Commerciali)*, 2001, *48*, pp. 133-43.

——, "Globalisation, International Investment and Stock Market Growth in Sub-Saharan Africa". Final research report submitted to the Institute of Developing Economies, JETRO, Japan, 2003.

BARRO R.J., "Inflation and Growth", *Federal Reserve Bank of St. Louis Review*, 1996, *78*, pp. 153-69.

BORSWORTH B.P. and COLLINS S.M., "Capital Flows to Developing Economies: Implications for Savings and Investment", *Brookings Papers on Economic Activity*, 1999, No. 1, pp. 143-80.

BRUNO M. and EASTERLY W., "Inflation Crises and the Long Run Growth", *Journal of Monetary Economics*, 1998, *41*, pp. 3-26.

DE MELLO L.R., Jr., "Foreign Direct Investment in Developing Countries and Growth: A Selective Survey", *Journal of Development Studies*, 1997, *34*, pp. 1-34.

——, "Foreign Direct Investment-Led Growth: Evidence from Time Series and Panel Data", *Oxford Economic Papers*, 1999, *51*, pp. 133-51.

FEDER G., "On Exports and Economic Growth", *Journal of Development Economics*, 1983, *12*, pp. 59-73.

FISHER S., "The Role of Macroeconomic Factors in Economic Growth", *Journal of Monetary Economics*, 1993, *32*, pp. 485-512.

FOSU A.K., "International Dimensions of African Economic Growth", CID Working Papers, Harvard University, 2000.

——, "International Dimensions of African Economic Growth", CID Working Papers, Harvard University, 2000.

HALL R. and JORGENSEN D.W., "Tax Policy and Investment Behaviour", *American Economic Review*, 1967, *57*, pp. 391-414.

ODEDOKUN M., "An Alternative Framework of Estimating Investment and Saving Functions for Developing Countries: An Application to Time Series Data for Sub-Saharan Africa Countries", *International Economic Journal*, 1992, *6*, pp. 49-73.

RAM R., "Exports and Economic Growth in Developing Countries: Evidence from Time Series and Cross Section Data", *Economic Development and Cultural Change*, 1987, *36*, pp. 51-72.

RAMIREZ M.D., "Foreign Direct Investment in Mexico: A Cointegration Analysis", *Journal of Development Studies*, 2000, *37*, pp. 138-62.

TYLER W.G., "Growth and Export Expansion in Developing Countries: Some Empirical Evidence", *Journal of Development Economics*, 1981, *9*, pp. 121-30.

ABSTRACT

The paper estimates the impact of foreign direct Investment (FDI) on capital accumulation, total factor productivity and economic growth. Panel data are provided for a sample of 12 African countries in the period 1990-2001. The results show that the stock of FDI has a positive impact on economic growth. In addition, exports and the stock of private investment have significant positive effects on growth. These results suggest the need to attract more foreign investment to Africa through appropriate fiscal monetary and institutional policies among others.

JEL classification: *HD2, EC9*
Keywords: *FDI, economic growth, Sub-Saharan Africa*

RISEC, Volume 50 (2003), No. 4, 581-591

THE IMPACT OF GUYANA'S FOREIGN DEBT BURDEN ON ITS RATE OF ECONOMIC GROWTH, 1971-2001

by
NELSON C. MODESTE* and HECTOR BUTTS**

1. *Introduction*

In 1971, Guyana's foreign debt[1] amounted to 55 percent of its GDP. Then, throughout the rest of the 1970's and 1980's, the foreign debt steadily rose until it reached 855 percent of GDP in 1991. Thereafter, with debt relief, the foreign debt burden fell to 182 percent of GDP in 2001. Accompanying this rise and fall in the foreign debt burden has been the contraction and expansion of the economy's real sector. Starting from the late 1960's and continuing into the early 1970's, output expanded moderately in the economy. Then, from the mid-1970's to the latter part of the 1980's, the economy contracted. Indeed between 1975 and 1988, output fell by more than fifty percent. With the adoption of economic reform measures in 1988,[2]

* Department of Economics and Finance, Tennessee State University, Nashville, TN (U.S.A.). E-mail: nmodeste@tnstate.edu

** Department of Business Administration, South Carolina State University, Orangeburg, SC (U.S.A.).

The authors would like to thank an anonymous referee for helpful comments on an earlier version of the paper.

[1] For this exercise, the foreign debt is gauged by the public and publicly guaranteed external debt of Guyana. This series is used because the data for this variable are consistently available for the period under review.

[2] Through the reform measures, Guyana obtained commitment of international support, which permitted it in 1989 to secure external debt relief, promote market-oriented practices, reduce state participation in the economy and encourage foreign investment. In particular, it benefited from balance of payments support in the form of payment support arrangement under

however, this downward trend in output was reversed. As such, for most of the 1990's, there was growth in the economy.

With studies by Kaminsky and Pereira (1996), Cunningham (1993), Borensztein (1990), Froot (1989), Sachs (1989), Berg and Sachs (1988) and others suggesting that the foreign debt burden adversely affects economic growth due to the debt overhang phenomenon,[3] the purpose of this study is to determine the extent to which economic growth in Guyana has been reduced by its foreign debt burden after controlling for the influence of other factors like exports, the investment-income ratio, and the growth in the population. To take into account the non-stationary nature of the time series variables in the empirical analysis, this study utilizes the co-integration and correction methodology.

Several factors make Guyana a very useful case for examining the effects of the foreign debt burden on economic performance. First, Guyana's foreign debt has been quite high with the stock of total outstanding public external debt as a percentage of income ranging from a low of about 52% in 1974 to a high of 855 percent in 1991. The seriousness of the country's foreign debt position is further affirmed by its designation as a heavily indebted poor country or HIPC. Second, under the auspices of the Paris Club and the HIPC initiative, Guyana has benefited from foreign debt relief. Countries providing such relief include the United States of America, Canada, and the United Kingdom, among other countries. Third, the data suggest that the decline and expansion of economic activity from the mid-1970's through the 1990's coincide with the growth and fall of the country's foreign debt position. Lastly, unlike the Cunningham (1993) study, which used pooled times series and cross-sectional data, an error correction framework is used to look at the relationship between the foreign debt burden and economic performance.

The rest of the paper is organized into three sections. Section 2 presents a model for studying the relationship between growth in income and growth in the foreign debt burden while taking into account other determinants of economic growth such as exports, the investment-income ratio and population growth. Section 3 describes the data used in the study as well as reports the results that have been generated by estimating the model from Section 2. In Section 4, the major points made in the paper are summarized.

---

the IMF and World Bank. The IMF provided extended arrangements, standby arrangements and ESAF, while the World Bank provided structural adjustment credit through the international Development Agency.

[3] For a contrasting view of the relationship between the foreign debt burden and the country's rate of economic growth, see Bulow and Rogoff (1990, p. 37).

2. *The Model*

The starting point for the empirical analysis of the relationship between the foreign debt burden and economic growth is the per capita version of the aggregate production function. In this form, per capita real income is determined by total factor productivity and the amount of capital individuals have to work with, that is:

$$(1) \qquad Y/N = A f(K/N)$$

where Y = real output; N = labor input; A = total factor productivity; and K = capital input. When equation (1) is converted into its growth form and allowance is made for total factor productivity to be influenced by export growth as well as by growth in the foreign debt burden, respectively, the following equation is obtained:

$$(2) \qquad (y - n) = \alpha_0 + \alpha_1 db + \alpha_2 x + \alpha_3 (k - n)$$

where y = growth in real output; n = growth in the population; $(y - n) = y'$ = growth in real per capita income or the growth in real output minus the growth in the population; db = growth in the foreign debt burden; x = growth in exports; k = growth in the capital stock; $(k - n)$ = growth in the per person capital stock. As data on the capital stock are unavailable, the variable k is measured by the ratio of investment to income, iy, or the investment rate. Consequently, when one substitutes the variable iy for the variable k in equation (2), one can now write the following estimable equation:

$$(3) \qquad (y - n) = \alpha_0 + \alpha_1 db + \alpha_2 x + \alpha_3 (iy - n) + u_t$$

where iy = ratio of investment to income or the investment rate; $(iy - n) = iy'$ = investment rate adjusted for growth in the population. The coefficients α_1, α_2, and α_3 measure the responsiveness of growth in real per capita income to the variables: db, x, and iy' respectively. Based on studies by Cunningham, Kaminsky and Pereira, Berg and Sachs, and others, the growth in real per capita income is expected to be negatively related to the growth in the foreign debt burden. For, as the foreign debt burden grows, there is the expectation that the productivity of labor and capital will decline. This outcome is anticipated given that an ever-increasing amount of the country's resources will have to go towards paying foreign creditors instead of supporting domestic investment activities. The growth in real per capita

income is meanwhile expected to be positively related to the growth in exports and the increase in the per person investment rate. In the case of the export variable, productivity is expected to rise as exports grow given the gains that accrue from comparative advantage and the easing of the foreign exchange constraint, among other things. The increase in the investment rate on a per person basis is meanwhile expected to lead to more growth as workers become more productive.

3. *Data and Empirical Analysis*

For this study, data for the period 1971 to 2001 were utilized. These data came from different sources. For exports, investment spending, income, and population, the data were taken from various issues of the International Monetary Fund, *International Financial Statistics Yearbook* (IFS), lines 90c, 93e, 99a, and 99z, respectively. For the consumer price index (CPI), the data for the period 1971 to 1990 were taken from the *IFS Yearbook*, line 64. For the remaining years, the CPI was constructed using the inflation rates as reported in various issues of the Inter-American Development Bank, *Economic and Social Progress in Latin America*. To measure the foreign debt burden, the ratio of the total outstanding public external debt to income was computed. The Bank of Guyana Half-Year Reports provided the information on the country's outstanding public external debt.

The behavior of the key variables in this study is depicted in Figures 1, 2, and 3, respectively. Figure 1 illustrates the performance of the economy through the variable on real income. As the trend line for that variable indicates, the economy grew moderately in the early 1970's. From 1975 until 1988, however, the economy was in a deep recession, with output contracting by more than 50 percent. A variety of external and internal factors, including declining terms of trade, the oil crises of the early and late 1970's, the nationalization of various industries in Guyana, and a failed approach to import-substitution industrialization have been identified as factors contributing to the recession. As the figure shows, there has been growth in the economy since the adoption of economic reform measures in 1988.

Figure 2 portrays the foreign debt burden as measured by the country's stock of foreign debt relative to income. Throughout the 1970's and 1980's, Guyana's foreign debt burden increased steadily until it reached a peak in 1991. Thereafter, with the provision of debt relief by international financial institutions (the Paris Club, the International Monetary Fund and the World Bank) and by countries such as Canada, the United States of America, and

the United Kingdom, among others, the foreign debt burden has fallen.

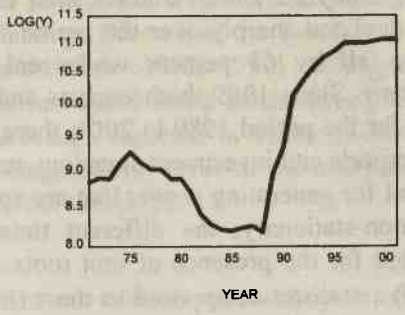

Figure 1. Real income (y)

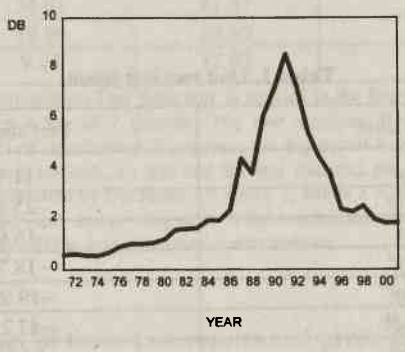

Figure 2. Foreign debt burden (db)

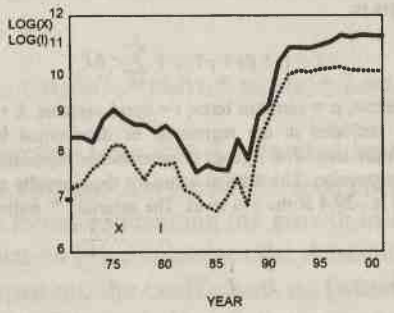

Figure 3. Real exports (x) and real investment (i)

Next, Figure 3 shows the behavior of investment spending and exports over the period being analyzed. Both variables, after exhibiting some growth in the early 1970's declined sharply over the period 1976 to 1988. For this period, real exports fell by 63 percent while real investment spending declined by 73 percent. Since 1989, both exports and investment spending have grown. In fact for the period 1989 to 2001, there has been a more than six fold increase in exports and investment spending, respectively.

Given the potential for generating results that are spurious if the variables in the model are non-stationary, the different time series variables for equation (3) are tested for the presence of unit roots. For this exercise, the Dickey-Fuller (1979) *z statistic* as opposed to the τ (*tau*) *statistic* is utilized for testing the unit root hypothesis. Since, according to Cook (2002), the *z* test statistic experiences less size distortions when the series contain structural breaks.[4] The results from these tests are reported in Table 1.

Table 1. Unit root test results

Variable	Test statistic (*z*)
y'	−11.99 (1)
db	−25.42 (1)
x	−16.08 (1)
iy'	−18.73 (1)
$\Delta y'$	−49.91* (0)
Δdb	−47.21* (0)
Δx	−53.44* (0)
$\Delta iy'$	−41.07* (0)

Notes: The statistic used for testing the null hypothesis of a unit root is $z = T(\hat{\rho} - 1)$. The test equation for this analysis is:

$$Y_t = \mu + \beta t + \rho Y_{t-1} + \sum_{i=1}^{k} c_i \Delta Y_{t-1}$$

where Y = series of interest; μ = constant term; t = trend variable; k = number of lagged first differenced regressors included in the regression as determined by Akaike Information Criterion, T = total sample size. The figures in parentheses represent the number of lagged terms included in the regression. The critical value for these results as taken from Davidson and Mackinnon (1993) is –29.4 at the 1% level. The asterisk, *, indicates significance at the 1% level.

[4] In the Cook paper, *K* is used as the symbol for the second unit root test statistic. In this paper, however, the symbol for that second unit root test is *z*. This change is made to avoid confusion with the *K* as defined in the production function.

They indicate that all of the variables are non-stationary in the form of y', db, x, and iy'. In light of this, each variable is expressed in first-difference form to achieve stationarity. Given that the four variables are integrated of order 1, the Johansen-Juselius (1990, 1992) co-integration test is used to determine if an equilibrium relationship exists among the variables: y', db, x, and iy'. Based on the results reported in Table 2, one must reject the null hypothesis of no co-integration in favor of the alternative hypothesis of co-integration among the four variables.

Table 2. Johansen's maximum eigenvalue test for cointegration

Likelihood ratio	5% critical value	1% critical value	Hypothesized No. of C.E.
62.27	53.12	60.16	None**
33.10	34.91	41.07	At most 1
10.58	19.96	24.60	At most 2
1.65	9.24	12.97	At most 3

Notes: The Johansen Co-integration Test procedure is applied to the four variable system (y', db, x, iy'). These results are for a lag of 2 periods. The two asterisks, **, denote rejection of the hypothesis at the 1% level of significance. To correct for degrees of freedom, the methodology in Cheung and Lai (1993) was utilized. To that end the test statistics produced by the econometric software program were multiplied by the factor, $(T - np)/T$, where T = sample size, n = number of variables in the cointegration test, and p = lag order in the cointegration test. The computer program utilized reports results only for the maximum eigenvalue statistic.

Finally, on the basis of these two results and given the Engle and Granger Representation Theorem (1987), which suggests that there is a correspondence between co-integration systems and error correction models (ECM)[5], the following restricted error correction model is utilized in this analysis:

$$\Delta y' = \alpha_0 + \alpha_1 \Delta db_t + \alpha_2 \Delta x_t + \alpha_3 \Delta iy' + \lambda ect_{t-1} + e_t \quad (4)$$

where ect = error correction term that is measured by the residuals from the estimation of equation (3); and Δ = first difference operator. Unlike equation (3), which depicts the forces influencing the growth in real per capita income over the long run, equation (4) emphasizes the dynamic aspect of the growth in income. For that equation, the coefficients α_i, (where i = 1, 2, 3) measure the short run response of growth in per capita income to changes in the

[5] For a discussion about this relationship, see Engle and Granger (1987).

variables: db, x, and iy'. The coefficient α_1 is expected to be negative. While the coefficients α_2 and α_3 should be positive. The coefficient for the error correction term, λ, is meanwhile expected to be negative.

Table 3. Estimation results for equation (4). Dependent variable: $\Delta Y'_t$

Regressor	Restricted ECM model	Unrestricted ECM model
Constant	–0.038 (–1.898)	–0.007 (–0.091)
Δdb_t	–0.157 (–1.702)	–0.226 (–2.601)
Δx_t	0.463 (12.596)	0.341 (7.586)
$\Delta iy'_t$	0.003 (0.012)	0.151 (0.456)
ect_{t-1}	–1.281 (–6.986)	–
y'_{t-1}	–	–1.275 (–8.315)
db_{t-1}	–	–0.398 (–3.483)
x_{t-1}	–	0.493 (4.936)
iy'_{t-1}	–	0.068 (0.284)
Dum	0.365 (4.322)	0.592 (6.354)
Adjusted R-squared	0.92	0.94
F-statistic	74.41	68.66
SER	0.104	0.086
JB statistic	0.94	0.69
LM (3)	5.470	4.604

Notes: The regression results are based on the period 1971-2001. SER is the standard error of regression; LM is the Lagrange Multiplier (Chi-squared) test for serial correlation; JB statistic is the Jarque-Bera test for normality of the residuals. The figures in parentheses alongside the estimated coefficients are the t-statistics.

Table 3 reports the results from estimating equation (4) by ordinary least squares. To control for the transitory effect of the government's Economic Recovery Program, a dummy variable with a value equal to 1 for 1989 –1991 and 0 otherwise was introduced into the estimated equation. The regression results for that model are quite interesting. First, they indicate that the estimated coefficient for the foreign debt burden variable, Δdb, not only possesses the correct theoretical sign but it is also statistically significant at the 5 percent level. In fact, the results suggest that a one percentage point increase in the country's foreign debt burden is associated with an almost two-tenths of a percentage point reduction in the growth of real per capita income. Second, the results indicate that the estimated coefficient for the export variable also has its correct theoretical sign and is statistically significant at the 1 percent level. For this variable, the results suggest that a

one percentage point increase in the growth of exports is associated with an almost four-tenths of a percentage point increase in the growth of real per capita income. Third, the estimated coefficient for the error correction term is statistically significant. This result would once again suggest that the variables in the equilibrium relationship are co-integrated. Additionally, the results indicate that as the growth in real per capita income diverged from its equilibrium value, there were, in the short run, corrective changes in the growth of real per capita income. Fourth, the statistical significance of the estimated coefficient for the dummy variable suggests that there was a positive transitory effect on the growth of real per capita income from the Economic Recovery Program. Finally, even though the estimated coefficient for the $\Delta iy'$ variable is statistically insignificant, the model, nevertheless, explains a significant proportion of the variation in the dependent variable given an adjusted R-squared value of 0.92. In addition, an F value of 74.41 also suggests that one must reject the null hypothesis that the estimated slope coefficients in the model are simultaneously equal to zero. Based on the Breusch-Godfrey serial correlation LM test, the null hypothesis of no autocorrelation is not rejected at the 1 percent level. Furthermore, the Jarque-Bera statistic suggests that one does not reject the null hypothesis that the residuals are normally distributed.

For comparison purposes, the following alternative unrestricted error correction model is also estimated:

$$(5) \qquad \Delta y_t' = \beta_0 + \beta_1 \Delta db + \beta_2 \Delta x + \beta_3 \Delta iy' + \beta_4 y'_{t-1} + \beta_5 db_{t-1} + \\ + \beta_6 x_{t-1} + \beta_7 iy'_{t-1} + e_t$$

This equation is obtained by using the lagged values of y', db, x, and iy' in the place of ect_{t-1} in equation (4). As reported in Table 3, the estimated coefficients for the dynamic variables are quite close to those reported for equation (4). Like the results for equation (4), the estimated coefficients for the variables Δdb, and Δx possess their correct theoretical sign and are statistically significant. The estimated coefficient for the other dynamic variable, $\Delta iy'$, remains statistically insignificant. With the exception of β_7, the estimated coefficients for the lagged values of the variables, y', db, and x are statistically significant. The estimated coefficient for the dummy variable is also statistically significant. From a policy perspective, these results suggest that reductions in Guyana's foreign debt burden should enhance the country's rate of economic growth. In addition, they suggest that even higher rates of economic growth are possible if the country's export sector can be expanded.

4. *Summary*

This paper looks at the impact of Guyana's foreign debt burden on its rate of economic growth. Within an error correction framework, a model that relates the growth in per capita output to the growth in the foreign debt burden, exports, and the investment rate adjusted for population growth is estimated. The results from this exercise show a strong negative relationship between Guyana's foreign debt burden and the performance of the economy. In light of this result, measures to reduce the country's stock of outstanding foreign debt are to be encouraged if faster growth in the economy is desired.

REFERENCES

BANK OF GUYANA, *Half-Year Report*, Georgetown, Guyana, (various issues).

BERG A. and SACHS J., "The Debt Crisis: Structural Explanations of Country Performance", *Journal of Development Economics*, 1988, *29*(3), pp. 271-306.

BORENSZTEIN E., "Debt Overhang, Credit Rationing and Investment", *Journal of Development Economics*, 1990, *32*(2), pp. 315-35.

BULOW J. and ROGOFF K., "Cleaning up Third World Debt Without Getting Taken to the Cleaners", *Journal of Economic Perspectives*, 1990, *4*(1), pp. 31-42.

CHEUNG Y.W. and LAI K.S., " Finite Sample Sizes of Johansen's Likelihood Ratio Tests for Co-integration", *Oxford Bulletin of Economics and Statistics*, 1993, *55*(3), pp. 313-28.

COOK S., "Spurious Rejection Using Dickey-Fuller Tests", *Applied Economic Letters*, 2002, *9*(9), pp. 557-62.

CUNNINGHAM R.T., "The Effects of Debt Burden on Economic Growth in Heavily Indebted Developing Nations", *Journal of Economic Development*, 1993, *18*(1), pp. 115-26.

DAVIDSON R. and MACKINNON J.G., *Estimation and Inference in Econometrics*, New York: Oxford University Press, 1993.

DICKEY D.A. and FULLER W.A., "Distribution of Estimators for Autoregressive Time Series with Unit Root", *Journal of American Statistical Association*, 1979, *74*(366), pp. 427-31.

ENGLE R.F. and GRANGER C.W.J., "Co-Integration and Error Correction Representation, Estimation, and Testing", *Econometrica* , 1987, *55*(2), pp. 251-76.

FROOT K.. "Buy Backs, Exit Bonds, and the Optimality of Debt and Liquidity Relief", *International Economic Review*, 1989, *30*(1), pp. 49-70.

INTER-AMERICAN DEVELOPMENT BANK, *Economic and Social Progress in Latin America*, Washington, DC., (various issues).

INTERNATIONAL MONETARY FUND, *International Financial Statistics (IFS) Yearbook*, Washington, DC., (various issues).

JOHANSEN S., "Statistical Analysis of Cointegration Vectors", *Journal of Economic Dynamics and Control*, 1988, *12*(2-3), pp. 231-54.

—— and JUSELIUS K., "Maximum Likelihood Estimation and Inference on Cointegration – with Applications to the Demand for Money", *Oxford Bulletin of Economics and Statistics* 1990, *52*(2), pp. 169-210.

—— and ——, "Testing Structural Hypothesis in a Multivariate Cointegration Analysis of the PPP and the UIP UK", *Journal of Econometrics*, 1992, *53*(1-3), pp. 211-44.

KAMINSKY G.L. and PEREIRA A., "The Debt Crisis: Lessons in the 1980's for the 1990's", *Journal of Development Economics*, 1996, *50*(1), pp. 1-24.

SACHS J.D., *Developing Country Debt and the World Economy*, Chicago: University of Chicago Press, 1989.

ABSTRACT

Within the context of the co-integration and error correction methodology, this paper examines the impact of Guyana's foreign debt burden on its rate of economic growth over the period 1971 to 2001. After allowing for growth in exports, investment-income ratio, and the growth in population, this study finds a significant negative relationship between growth in Guyana's foreign debt burden and growth in real income in the economy. Given this, debt relief is seen as being highly desirable for the promotion of growth in the Guyanese economy.

JEL classification: *O11, O54, F34*
Keywords: *foreign debt burden, economic growth, debt relief*