International Lending By US Banks

by

John M. Barron
Department of Economics
Purdue University
West Lafayette, IN 47907
(765) 494-4451
e-mail: barron@mgmt.purdue.edu

Neven T. Valev
Department of Economics
Georgia State University
Atlanta, GA 47907
(404) 651-0418
e-mail: prcntv@langate.gsu.edu

Abstract

This paper develops a portfolio choice theory in which the sequential order of decision makers is endogenously determined. We examine the decision of investors to purchase current information on returns to particular projects or to wait and later base their actions on information inferred from the decisions of others. Because the quality of that inference increases in the number of investors who purchase information, some investors will find it advantageous to wait and invest later. Thus investors endogenously separate into leaders and followers. We show that the extent of such separation is related to the persistence in states of the world across time. Some implications of the model are tested using data on international lending by US banks for the 1982-1994 period. The empirical results support the main predictions of the model.

JEL Classification: D82 Asymmetric and Private Information; F24 International Lending and Debt Problems

The authors are grateful to John A. Carlson, Kenneth Matheny, the participants at the Midwest Economics Association Meeting in Kansas City, MO, March, 1997, and two referees for valuable comments and suggestions.
1. Introduction

A common phenomenon is that the decisions of some individuals are influenced by what others have already done. From a Bayesian perspective, it is not surprising that the actions of those who go first, the leaders, are often mimicked by those who follow as the leaders’ actions can provide information that influences the choices made by later decision makers. Two questions arise in such sequential decision-making settings: Can we determine *ex ante* who will be the leaders and who will be the followers? Can we identify observable factors that determine the extent to which the followers’ actions mimic the leaders’ actions?

The purpose of this paper is to answer the above two questions in a model that highlights how heterogeneity across agents can result in an endogenous determination of who leads and who follows. Specifically, we examine in the context of an investment decision how wealth differences influence investors’ choices of either purchasing current information or waiting and later basing their actions on information inferred from the decisions of others. Because the quality of that inference increases in the number of investors who purchase information, those with low wealth will find it advantageous to wait and invest later. Thus investors endogenously separate into leaders and followers. We show that the extent of such separation is related to the persistence in states of the world across time. In particular, if states are not correlated across periods, prior investment behavior provides no valuable information about the state in the current period. Thus, fewer investors will find it optimal to wait and follow others.

Some of the implications of the model are tested using data on international lending by US banks for the 1982-1994 period.\footnote{We pick international bank lending primarily because it is populated by decision makers that differ substantially in their size, which the theory suggests is one important determinate of the value of acquiring private information, and thus who leads and who follows.} The empirical results support the main predictions of the model. First, we
find that small banks follow the lending behavior of large banks with regard to which countries to lend to. The theory developed also suggests that the extent of such behavior will depend on the persistence in states of the world. We develop a measure of persistence in terms of economic conditions and find that, as predicted, small banks follow large banks to a greater extent in countries where there is more persistence in economic conditions.

The analysis in this paper builds upon the work of Jain and Gupta (1987), who provide some evidence of “herd behavior” in international lending. Our analysis differs from theirs in two respects. First, with regard to the theoretical analysis, we construct a model that highlights the important role of persistence in determining the extent to which small banks follow large. In doing so, we incorporate some of the features of what has been termed the “herd behavior” phenomenon. In herd behavior models along the lines of Banerjee (1992), the choice by some to follow is “socially excessive” in that such choices hide useful information from others. In our model, the choice by some to follow means not purchasing information that others would find useful. Second, we extend the empirical analysis to a

---

2 For instance, Banerjee (1992) states that “…the very act of trying to use the information contained in the decisions made by others makes each person’s decision less responsive to her own private information and hence less informative to others. Indeed, we find that in equilibrium the reduction in informativeness may be so severe that in an ex ante welfare sense society may actually be better off constraining some of the people to use only their own information.” This type of “herd behavior” can be accentuated if the future payoff of the decision maker depends on whether her decision differs from the decisions of others. Sharfstein and Stein (1990) develop a simple example of just this “agency/reputation” herd behavior. In their model, investment managers always mimic the decisions of others to enhance their reputation. If managers were instead to base their decisions on private information and Bayesian updating alone, the possibility of behavior different from other managers would increase the likelihood of being identified as incompetent, and thus receiving lower future wages. Calvo and Mendoza (1996, 1997) study the effect of globalization on the extent of herd behavior in international markets. They develop a model of international portfolio diversification and show that increasing the number of available options for international investment reduces the incentive of individual investors to acquire costly country-specific information. In particular, “highly diversified investors have lower incentives to acquire information than investors with fewer investment opportunities. This in turn results from the fact that as the number of countries in which to invest increases, the marginal gain from information gathering eventually declines.” As a result, “rational investors become extremely susceptible to “small” news.”
broader time period and a larger number of countries, and we provide new tests of the role of persistence in explaining the patterns of international lending.

The next section presents a model of an investor facing incomplete information. We then discuss the empirical implications with respect to international lending. In the final two sections of the paper, we discuss the data, present the empirical tests of the theory, and offer concluding remarks.

2. A Simple Model

We begin our analysis with a simple portfolio problem for an agent $i$ with wealth to invest $W_i$. There are two types of investments: a risky asset and a risk-free asset. The risky asset pays one of two possible returns: $R_H$ or $R_L$. We say that the state $S_i = H$ has occurred in period $t$ when the return is $R_H$ and that the state $S_i = L$ has occurred if the return is $R_L$. Thus, if the decision is to invest in the risky asset, investor $i$'s investment yields $R_{W_i}$ if $S_i = H$ and $R_{L_i}$ if $S_i = L$. If the investor chooses the risk-free asset, the gross return $R_0$ is such that $R_H > R_0 > R_L$ and the investor obtains $R_{W_i}$. Investors differ in wealth, with $W_i \in (0, \bar{W}]$, $i = 1, \ldots, n$. The prior probability that the state $H$ will occur in period $t$ is $\alpha_t$.

2.1 The Decision to Become Informed

Each period $t$ investor $i$ has the option of purchasing a signal $s = \{H, L\}$ at cost $C$. The signal is informative as to the state in that:

$$P(s_i = H|S_i = H) = P(s_i = L|S_i = L) = q$$

$q \in (\frac{1}{2}, 1)$

If an investor purchases the signal $s$, she invests in the risky asset in period $t$ only if $s_i = H$. Below we determine the separation of the $n$ agents into those who purchase information and those who do not. We call individuals in the first group the informed, and those in the second group the uninformed.
Agent $i$ at the start of period $t$ has three options: first, to not purchase the signal $s_t$ and to not invest in the risky asset, in which case investor $i$ obtains expected utility:

\[(2) \quad U_N = R_0 W_i \quad .\]

The second option is to not purchase the signal $s_t$ and to invest $W_i$ in the risky asset. In that case, the investor obtains expected utility:

\[(2') \quad U_I = [\alpha_i R_H + (1 - \alpha_i) R_L] W_i \quad .\]

Third, investor $i$ may purchase the signal $s_t$ and invest $W_i$ in the risky asset if $s_t = H$. That option yields expected utility:

\[(2'') \quad U_P = [\alpha_i R_H + (1 - \alpha_i) R_0 - \alpha_i (1 - q)(R_H - R_0) - (1 - \alpha_i)(1 - q)(R_0 - R_L)] W_i - C \quad .\]

Expected utility \((2'')\) is written in terms of type I and type II errors. In particular, the first two terms in the brackets represent the expected utility if investor $i$ correctly recognizes the true state of the world. The third and fourth terms are the losses associated with type I and type II errors, i.e. not recognizing a good state and not recognizing a bad state, respectively.

Investor $i$ chooses among the three options \{\(U_N, U_I, U_P\)\} the one that yields the greatest expected return. The expected return is thus

\[(3) \quad V_i(\alpha_i) = \max \{U_N, U_I, U_P\} \quad .\]

The choice will depend on the prior probability of a good state, $\alpha_i$, and on the level of wealth, $W_i$. Note that at $\alpha_i = 0$,

\[U_N = R_0 W_i \quad > \quad U_P = [R_L + q (R_0 - R_L)] W_i - C \quad > \quad U_I = R_L W_i \quad ,\]
indicating that the best option for the individual is to not invest given a sufficiently low probability of a good state, while the worst option is to invest. However, with an increase in the prior probability of a good state, note that:

\[
\frac{\partial U_N}{\partial \alpha_i} = 0 ,
\]

\[
\frac{\partial U_p}{\partial \alpha_i} = [q(R_H - R_0) + (1-q)(R_0 - R_L)] W_i , 
\]

\[
\frac{\partial U_i}{\partial \alpha_i} = [R_H - R_L] W_i ,
\]

so that

\[
\frac{\partial U_N}{\partial \alpha_i} < \frac{\partial U_p}{\partial \alpha_i} < \frac{\partial U_i}{\partial \alpha_i} .
\]

Thus, an increase in the prior probability of a good state increases the gain to simply investing at a greater rate than the increase in the gain to purchasing an additional signal prior to making the investment decision. At \( \alpha_i = 1 \),

\[
U_N = R_0 W_i < U_p = [R_H + q(R_H - R_0)] W_i - C < U_i = R_H W_i ,
\]

indicating that if the probability of a good state is sufficiently high, the best option is to simply invest.

The gain to purchasing an informative signal rises as \( \alpha_i \) approaches mid-range values. In particular, a signal will be purchased when \( \alpha_i \) falls in the range \((\alpha_{HL}, \alpha_H)\) where \( \alpha_L \) solves \( U_N = U_p \)

and \( \alpha_H \) solves \( U_p = U_i \). We thus define \( \alpha_L \) and \( \alpha_H \) as follows:

\[
\alpha_L = \frac{(1-q)(R_0 - R_L) + C/W_i}{q(R_H - R_0) + (1-q)(R_0 - R_L)} = \alpha_L (C/W_i)
\]

and

\[
\alpha_H = \frac{q(R_0 - R_L) - C/W_i}{(1-q)(R_H - R_0) + q(R_0 - R_L)} = \alpha_H (C/W_i)
\]
Figure 1 illustrates an investor with $\alpha_L < \alpha_H$. For this investor, all three options (invest, no signal; do not invest, no signal; purchase a signal and invest based on the signal) are viable depending on the prior probability of the good state. Note that the line denoting $U_N$ (the value of the no invest option) identifies this option as having the greatest value for low $\alpha$, but having a value that is unaffected by changes in $\alpha$. On the other hand, the line denoting $U_I$ (the value of the invest option) identifies this option as having the lowest value of the three options for low $\alpha$. But the value of this option increases more rapidly than the value of either alternative option, in particular more rapidly than the value to purchasing a signal prior to making an investment decision ($U_p$), with increases in $\alpha$. Thus, for very large and very small $\alpha$, there is no gain to purchasing information and investor $i$ purchases information if $\alpha_i \in (\alpha_L, \alpha_H)$.

**Figure 1**

*Optimal Investment Strategy for a Wealthy Agent*
The range of $\alpha_i$ that will induce an individual to purchase the signal prior to the investment decision is increasing in wealth and decreasing in the cost of the signal. These results follow directly from equations (4) and (5), which indicate that $\alpha_L$ depends inversely and $\alpha_H$ depends directly on the wealth-to-signal-cost ratio $W_i / C$. Figure 2 illustrates this relationship between the wealth-to-signal-cost ratio and the range of prior probabilities of a good state for which it is the best option to purchase information prior to the investment decision.

Figure 2
Range of Priors for Signal Purchase By Wealth-to-Signal-Cost Ratio

![Graph showing the range of priors for signal purchase by wealth-to-signal-cost ratio]

Note that for a sufficiently low wealth or sufficiently high signal cost, there will be no prior probability of a good state at which information will be purchased. The critical wealth-to-cost level below which investor $i$ will never purchase information, $W_p$, solves $\alpha_L = \alpha_H$, and is given by:

$$W_p = \frac{C(R_H - R_L)}{(2q-1)(R_H - R_b)(R_b - R_L)}$$
We have established that investor $i$ will purchase information in period $t$ if she has sufficient wealth relative to the cost of the signal and if the probability of a good state $\alpha_i$ is not too high or too low. Let $H(W)$ denote the underlying distribution of wealth across $n$ agents. The maximum number of investors that become informed occurs at the prior probability of a good state $\alpha_p = \alpha_H(C/W_p) = \alpha_L(C/W_p)$, and is given by $(1 - H(W_p))n$. For $\alpha_i < \alpha_p$, the number of individuals who become informed is defined by $\alpha_L(C/W_i) < \alpha_i$, and thus from (4) is given by $n'_i = (1 - H(W_i^L))n$ where from equation (4):\(^3\)

\begin{equation}
W_i^L = C / [-(1-q)(1-\alpha_i)(R_0 - R_L) + \alpha_i q(R_H - R_L)] < W_p
\end{equation}

For $\alpha_i > \alpha_p$, the number of individuals who become informed is defined by $\alpha_H(C/W_i) > \alpha_i$, and thus from (5) is given by $n'_i = (1 - H(W_i^H))n$ where from equation (5):

\begin{equation}
W_i^H = C / [q(1-\alpha_i)(R_0 - R_L) - \alpha_i (1-q)(R_H - R_L)] < W_p
\end{equation}

The above results indicate that the number of investors who become informed depends on the informativeness of the signal ($q$) and the prior probability of a good state ($\alpha_i$). A more informative signal or a prior probability of a good state that lies further toward the extremes will lower $W_i^H$ and $W_i^L$, and thus increase the proportion of individuals who will become informed.

The next step in our analysis is to consider the determinants of $\alpha_i$, the prior concerning the probability of a good state. In doing so, we establish a link between the number and actions of the informed in one period and the desire to become informed in the next. This leads in subsequent sections to a discussion of the equilibrium distribution of priors that will emerge. This distribution determines, among other factors, the expected number of individuals in any given period who will be informed (the

---

\(^3\) Equation (7) is derived by solving (4) for wealth, then substituting the prior probability of a good state ($\alpha_i$).
“leaders”) and the number whose actions will be dictated by the actions of the informed in the prior period (the “followers”).

2.2 Transmission of Information Across Periods

To introduce structure to the transition across periods between states, let us assume that states of the world follow a Markov process with the following transitional probabilities:

\[
\begin{align*}
P(S_t = H | S_{t-1} = H) &= \beta \\
P(S_t = L | S_{t-1} = H) &= 1 - \beta \\
P(S_t = L | S_{t-1} = L) &= \varphi \\
P(S_t = H | S_{t-1} = L) &= 1 - \varphi
\end{align*}
\]

so that

\[
\alpha_t = \beta \alpha_{t-1} + (1 - \varphi)(1 - \alpha_{t-1}).
\]

Equation (10) implies that \( \alpha_t \in [1 - \varphi, \beta] \). In what follows, we assume for presentation purposes that \( \beta = \varphi \).

We assume that investors gain information on the state of the world by observing the behavior of the informed investors in the preceding period. Of the \( n_{t-1}^I \) number of informed (purchased-a-signal) investors in period \( t-1 \), let the number of informed investors who invested in the risky asset in period \( t-1 \) be \( j_{t-1} \). The proportion of informed investors who invest in the risky asset is informative since informed investors invest in the risky asset only if \( s_{t-1} = H \). In particular, the number of informed investors who invest in the risky asset in period \( t-1 \) given that \( S_{t-1} = H \) has a binomial distribution with parameters

---

4 Such a Markov-switching process is analyzed by Hamilton (1989).

5 This assumption does not qualitatively affect any of the results.

6 Note that investors do not observe the true realization of the state for the prior period. One can interpret this assumption as follows: although each investor knows what his individual project paid at the end of last period, this does not help him locate a good project this period. There is a “fresh” supply of new projects each period and the only way in which they are related between periods is through their expected quality.
\((n_{t-1}^I, q)\), and the number of informed investors who invest in the risky asset in period \(t-1\) given that \(S_{t-1} = L\) has a binomial distribution with parameters \((n_{t-1}^I, 1-q)\).

At the start of each period, the observed number of previously informed agents who invested acts as a signal concerning the prior state. In particular, letting \(j^*\) denote the realized number of the \(n_{t-1}^I\) informed investors who invest in the risky asset in period \(t-1\), we have that:

\[
P(j_{t-1} = j^* \mid S_t = H) = \binom{n_{t-1}^I}{j^*} q^{j^*} (1-q)^{n_{t-1}^I - j^*} \quad \text{and} \quad P(j_{t-1} = j^* \mid S_t = L) = \binom{n_{t-1}^I}{j^*} (1-q)^{j^*} q^{n_{t-1}^I - j^*}
\]

Thus the updated probability that a good state existed for period \(t-1\) can be expressed as:

\[
P(S_{t-1} = H \mid \alpha_{t-1}, j^*, n_{t-1}^I, q) = \frac{\alpha_{t-1} q^{j^*} (1-q)^{n_{t-1}^I - j^*}}{\alpha_{t-1} q^{j^*} (1-q)^{n_{t-1}^I - j^*} + \alpha_{t-1} q^{n_{t-1}^I - j^*} (1-q)^{j^*}} = \alpha_{t-1}^*
\]

There are two main implications of (11). First, for a given number of informed investors in the preceding period, \(n_{t-1}^I\), the updated probability of a good state in the prior period, \(\alpha_{t-1}^*\), increases in \(j^*\):

\[
\frac{\partial \alpha_{t-1}^*}{\partial (j^*/n_{t-1}^I)} \geq 0
\]

That is, for a given number of informed investors, a greater proportion of those investors that invest in the risky asset in period \(t-1\) indicates a greater probability of a good state in period \(t-1\).

Second, keeping the proportion that invest, \(j^*/n_{t-1}^I\), constant and differentiating \(\alpha_{t-1}^*\) with respect to the number of informed investors, \(n_{t-1}^I\), yields:

\[
\frac{\partial \alpha_{t-1}^*}{\partial n_{t-1}^I} \begin{cases} 0 & \text{if } (j^*/n_{t-1}^I) > 1/2 \\ \leq 0 & \text{if } (j^*/n_{t-1}^I) \leq 1/2 \end{cases}
\]

In the limit, as \(n_{t-1}^I \to \infty\), \(\alpha_{t-1}^* \to 1\) if \(j^*/n_{t-1}^I > 1/2\) and \(\alpha_{t-1}^* \to 0\) if \(j^*/n_{t-1}^I < 1/2\). That is, for a given proportion of informed investors who invest, a greater number of informed investors improves the
quality of the inference. Thus, an agent who decides to become informed in period \( t-1 \) improves the quality of information inherited by all agents concerning the state in period \( t-1 \).

Given the updated prior \( \alpha^*_{t-1} \) of a good state in period \( t-1 \) and the transition probability \( \beta \) for not changing states, the prior probability for the good state in period \( t \), \( \alpha_t \), is given by:

\[
(14) \quad \alpha_t = \beta \alpha^*_{t-1} + (1-\beta)(1-\alpha^*_{t-1})
\]

Equation (14) indicates that a greater probability of a good state in period \( t-1 \), \( \alpha^*_{t-1} \), implies a greater probability of a good state in period \( t \) if \( \beta > \frac{1}{2} \)

**2.3 The Equilibrium Distribution of \( \alpha \).**

In section 2.1, we showed that \( \alpha_t \), the probability of a good state in period \( t \), affects the proportion of agents who find it optimal to become informed in a given period. In section 2.2, we linked this perceived probability of a good state to the prior period’s investment behavior of the informed. The strength of that link depended on the number of informed investors in the prior period, the quality \( q \) of the signal \( s \), and on the underlying transition probability \( \beta \) that defines the persistence of a particular state. In this setting, we seek to answer the following question. How will a change in persistence of states or in the quality of the signal \( s \) affect the extent to which individuals become informed as opposed to basing investment decisions on the actions of the informed in the prior period?

To answer the above question requires a characterization of the equilibrium distribution of the perceived probabilities of success, that is the likelihood of various probabilities of a good state \( \alpha_t \) for a randomly chosen period \( t \). This equilibrium distribution of \( \alpha_t \) denoted by \( G(\alpha) \), will be the limiting one as the system evolves over time from any initial probability of a good state. From (4) and (5) and the underlying distribution of wealth across agents \( H(W) \), it follows that the expected proportion of
individuals who become informed is \[ \int G(\alpha_L) - G(\alpha_H) \, dH(W). \] Below we explore in some detail the nature of the equilibrium distribution of updated priors \( G(\alpha) \). In addition, we consider how changes in the persistence parameter \( \beta \) and quality of the signal parameter \( q \) affect this distribution of probabilities of a good state, and the resulting implications concerning the expected proportion of agents who become informed.

To gain some intuition with regard to the character of the equilibrium distribution of priors concerning the probability of a good state, consider two illustrative extreme cases. In the first case, the state of the prior period is known with certainty and the state of the world in the prior period is correlated with the state in the current period (i.e. \( \beta > \frac{1}{2} \) and \( 1 - \varphi < \frac{1}{2} \)) Then from (10) \( \alpha = \{ \beta, 1 - \varphi \} \) with \( \alpha = \beta \) if the prior period’s state is \( S = H \) and \( \alpha = 1 - \varphi \) if the prior period’s state is \( S = L \). In the second case, the state of the world in the prior period is not correlated with the current period’s state (i.e., \( \beta = \frac{1}{2} \) and \( 1 - \varphi = \frac{1}{2} \)) Then from (10) \( \alpha = \frac{1}{2} \)

The case we are interested in falls between the above two extreme cases. In this intermediate case, unlike the first extreme case, after each switch of the state variable from \( H \) to \( L \) or from \( L \) to \( H \), a learning process begins that only gradually reveals the true underlying state of the world. And, unlike the second extreme case, learning the state of the world for the prior period provides some information about the current period’s state given that states are correlated.

With \( \beta = \varphi \), the expected duration of each state is \( \beta / 1 - \beta \). Consider a switch from \( L \) to \( H \) and let \( \beta / 1 - \beta = T \) periods. Then after the switch in period \( t \), the expected path of \( \alpha \) is given by

---

7 In this case, an \( H \) period is expected to last for \( \beta - 1 - \beta \) periods and an \( L \) period is expected to last for \( \varphi / 1 - \varphi \) periods (the expected duration of a state is a geometric random variable).

8 If \( \beta = \varphi \) the results are qualitatively the same. In what follows, we compare two processes with high and low \( q \). Dealing with different persistence for \( H \) and \( L \) requires that we keep track of two cases that are qualitatively...
\[ \alpha_D + (E_i \alpha_{t+1} - \alpha_D) + (E_i \alpha_{t+2} - E_i \alpha_{t+1}) + \ldots + (\alpha_U - E_i \alpha_{t+T-1}) = \alpha_U, \]

where \( \alpha_D \) is the initial value of \( \alpha \) and \( \alpha_U \) is the value to which \( \alpha \) converges before the next switch (from \( H \) to \( L \)). Note that, with \( \beta > \frac{1}{2} \) and \( S = H \):

\[ E_{t-1}(\alpha | S_{t-1} = H, \alpha_{t-1}) - \alpha_{t-1} > 0 \]

and, from (11) and (14), with \( \beta > \frac{1}{2} \):

\[ \frac{\partial}{\partial q} [(E_{t-1}\alpha | S_{t-1} = H, \alpha_{t-1}) - \alpha_{t-1}] > 0 \]

Similarly, after a switch from \( H \) to \( L \), the expected dynamics of \( \alpha \) is given by:

\[ \alpha_U - [\alpha_U - E_i \alpha_{t+1}] - [E_i \alpha_{t+2} - E_i \alpha_{t+1}] - \ldots - [E_i \alpha_{t+T-1} - \alpha_D] = \alpha_D \]

so that, with \( \beta > \frac{1}{2} \) and \( S = L \):

\[ E_{t-1}(\alpha | S_{t-1} = L, \alpha_{t-1}) - \alpha_{t-1} < 0 \]

and, from (11) and (14), with \( \beta > \frac{1}{2} \):

\[ \frac{\partial}{\partial q} [(E_{t-1}\alpha | S_{t-1} = L, \alpha_{t-1}) - \alpha_{t-1}] < 0 \]

In other words, after a switch from \( L \) to \( H \), \( \alpha \) is expected to converge from a lower to a higher value, and after a switch from \( H \) to \( L \), \( \alpha \) is expected to converge from a higher to a lower value. In expected terms, with \( \beta = \varphi \), \( H \) and \( L \) states will be recognized with the same precision, i.e. \( \alpha_U = 1 - \alpha_D \). To see this, recall that, with \( \beta = \varphi \), the expected duration of each state is \( \beta / (1 - \beta) \) and, equivalent.

\[ E_{t-1}\alpha = \sum_{j=0}^{n-1} \alpha_j(j_{t-1})f(j_{t-1}), \] where \( \alpha(j_{t-1}) \) is given by (14) and \( f(j) = b(q, n_t) \) if \( S = H \), and \( f(j) = b(1-q, n_t) \) if \( S = L \).
also, that the processes of recognizing $H$ and $L$ states are equivalent. Hence, for given $q$ and $\beta$, $\alpha$ is distributed in $[\alpha_D, \alpha_U]$ with $\alpha_U = 1 - \alpha_D$. Finally, from (17) and (17'), we can establish that:

\begin{equation}
\frac{\partial \alpha_U}{\partial q} > 0
\end{equation}

i.e. states will be recognized faster and with greater precision if the signal $s$ is better quality.

Let $\alpha_D < \alpha_H < \alpha_L < \alpha_U$ ($\alpha_H$ and $\alpha_L$ are given by (6) and (7)). Let $t^*$ denote the expected number of periods between two switches during which $\alpha$ is in the interval $(\alpha_L, \alpha_H)$ as opposed to $[\alpha_D, \alpha_L] \cup [\alpha_L, \alpha_U]$. Then $t^*/T$ can be interpreted as the probability that $\alpha \in (\alpha_L, \alpha_H)$. We are interested in the effect of quality of information $q$, and persistence $\beta$ on $t^*/T$.

An example will help clarify our discussion. In Figure 3 we plot the path of $\alpha$ for two examples in which persistence $\beta$ and the number of informed investors are the same. In period 1, the state switches
from $L$ to $H$ and, in period 10, the state switches from $H$ to $L$ so that both states last for 9 periods.\(^\text{10}\)

The two examples differ in the quality of the signal $s$, $q$. In particular, the path with greater amplitude is given by a greater value for $q$. Recall that what we are interested in is the proportion of periods in which $\alpha$ will have values outside of the interval $(\alpha_L, \alpha_H)$. For illustration, we chose $\alpha_H = .6$ and $\alpha_L = .4$. Clearly, with greater $q$, $\alpha$ “spends more time” outside $(\alpha_L, \alpha_H)$. In the proof of Proposition 1, we establish that this result will hold for any $\alpha_H$ and $\alpha_L$ in $(\alpha_D, \alpha_U)$.

**Proposition 1:** Let $\alpha_D < \alpha_H < \alpha_L < \alpha_U$. Then $\Pr(\alpha_L < \alpha < \alpha_H)$ decreases in the quality of the signal $s$, $q$.\(^\text{11}\)

Next, we want to determine the effect of greater persistence $\beta$ on the probability that $\alpha \in (\alpha_H, \alpha_L)$. The effect of $\beta$ on that probability is very similar to that of quality of the signal. With greater persistence $\beta$, true states are recognized more precisely.\(^\text{12}\) In addition, greater persistence $\beta$ is consistent not only with greater amplitude of the dynamic path but also with lower frequency of switches between states (recall that the expected duration in a state is $\beta/1-\beta$). This second effect, however, only magnifies the first with respect to the probability that $\alpha$ lies between $\alpha_L$ and $\alpha_H$.

To illustrate, in Figure 4 we show two examples with the same quality of information $q$, and number of informed investors. The two examples differ in persistence $\beta$. In particular, the frequency of switches for the path with greater amplitude is 9 periods ($\beta = .9$) and the frequency of switches for the other path is 3 periods ($\beta = .75$). With $\alpha_H = .6$ and $\alpha_L = .4$, $\alpha$ clearly has more values outside $(\alpha_L, \alpha_H)$ when $\beta = .9$. We thus have:

\(^\text{10}\) An expected duration of 9 corresponds to $\beta = 0.9$.
\(^\text{11}\) The proof is in Appendix A.
\(^\text{12}\) From (14), $\partial \alpha / \partial \beta > 0$ if $\alpha^* > 1/2$. Hence, the slope of the path of $\alpha$ is greater under high $\beta$. 
Proposition 2: Let $\alpha_D < \alpha_H < \alpha_L < \alpha_U$. Then, $\Pr(\alpha_L < \alpha < \alpha_U)$ decreases in persistence $\beta$.

In the proof of Proposition 2, we establish that this result will hold for any $\alpha_H$ and $\alpha_L$ in $(\alpha_D, \alpha_U)$.\(^\text{13}\)

Finally, the effect of $\beta$ and $q$ on the proportion of agents who purchase information is a straightforward result of Propositions 1 and 2. Recall, that agents with sufficient wealth purchase information if $\alpha \in (\alpha_H, \alpha_L)$. Thus,

**Proposition 3.** The average proportion of agents who purchase information decreases in persistence $\beta$ and in the quality of the signal $s$, $q$.

2.4. Some welfare implications.

Greater $\beta$ or/and $q$ shifts $\alpha$ in probabilistic terms from the midpoints to the tail points, and thus, by Blackwell’s (1961) theorem, increases the expected utility of individual agents. If social welfare is

---

\(^{13}\) The proof of Proposition 2 is in Appendix A.
defined as the sum of expected utility across agents, then greater persistence, \( \beta \), and/or quality of information, \( q \), are welfare improving. Intuitively, with greater \( q \) and \( \beta \), the return on purchased information is greater as its benefits extend for more than one period. Thus, the same overall quality of information can be achieved at a lower cost as inheriting better information implies that some agents will not find it optimal to purchase information. A key point here is that agents ignore the future benefits for others of the information that they purchase in the current period. Hence, given \( \beta \) and \( q \), social welfare will increase if the marginal investor were to purchase information.

Greater persistence and/or quality of information, however, while welfare-improving also provide for some interesting short-term dynamics. Specifically, greater persistence or quality of signal means an increase in the proportion of agents who follow (do not become informed) so that after each switch in the state variable, there will be a greater number of investors who act on information that is “too old”. Thus, inducing the marginal investor to purchase information not only enhances long-term social welfare, but also decreases the proportion of agents who will continue investing in the risky asset even after the state has switched from \( H \) to \( L \) and visa versa.

2.5. Overlapping generations.

A natural extension of the above set-up is to allow for agents who live for two periods and can purchase information in the beginning of both periods. In any given period, old agents will act as the agents we have in the discussion so far. Young agents, however, will consider the effect of an additional signal on the information they will inherit when they are old, i.e. young agents will take into consideration the future externality of current information. In this sense, the proportion of young agents who purchase information will be greater, and thus the young are more likely than the old to be leaders rather than followers.
3. Empirical Tests of the Model

Our analysis in the prior section suggests that one important distinguishing feature of investors that determines whether they are leaders or followers is the level of their wealth. In the context of international bank lending, we can use bank size ("large" banks versus "small" banks) to differentiate investors on the basis of their wealth. For international lending, it is natural to identify short-term loans to a foreign country as investments in what we have termed the risky asset. The signal concerning the risky investment would thus be the acquisition of detailed, country-specific knowledge of a particular country’s current borrowers.

We assume that large U.S. banks hold sufficient wealth so that they always choose to purchase information on foreign borrowers, while small banks may or may not purchase current information depending on the persistence in states of the world across time in that country. By Proposition 3, it will be less likely that small banks find it optimal to purchase information in countries with greater persistence in states of the world. Thus, in countries exhibiting greater persistence, small banks respond to changes in states one period later than large banks. In this sense, the above discussion suggests the following testable hypotheses.

1. An increase (decrease) in the level of investment in a country by large banks in one period will lead to an increase (decrease) in the level of investment by small banks in the subsequent period. However,

2. The extent of such behavior should depend on the extent of persistence across time in states, as indicated by the magnitude of $\beta$.

In particular, the larger the $\beta$, the stronger will be the correlation between changes in the level of investment by large banks and subsequent changes in investment levels by small banks.
To test the predictions of the model, we start with the semiannual data on U.S. bank outstanding foreign claims provided by the “Country Exposure Lending Survey” published by the Financial Institutions Examination Council at the Board of Governors of the Federal Reserve.\textsuperscript{14} The results of the survey originated in 1975 as an annual publication. Beginning in 1977, the survey results were published semiannually. But it was not until 1982 that data breaking down bank loans by bank size were made public.\textsuperscript{15} Our analysis is thus restricted to the period 1982 - 1994.

Foreign lending activity is broken down in several ways. First, three broad classes of U.S. banks - large money center banks, medium-sized banks, and small regional banks – are identified. Contained within these three groups are all US banks with either a foreign branch or at least $20 million in outstanding foreign loans. We exclude medium-sized banks from the analysis below in order to focus on a clear polarization in terms of types of banks.\textsuperscript{16} Besides bank size, information on international lending by banks is broken down by various maturities. In the analysis to follow, we focus on short-term (less than a year) debt. The reason for this is that the theory considers only short-term lending. However, for completeness we also provide estimates using measures of all lending activity.

Our analysis is restricted to 40 countries.\textsuperscript{17} We exclude countries where small banks do not have sufficient presence (defined as at least USD50 million in outstanding loans).\textsuperscript{18} A list of the countries in

\begin{itemize}
\item[\textsuperscript{14}] Outstanding claims are defined as the US dollar amount of disbursed debt held by US banks on a given date (December 31st and July 31st). These are amounts owed US banks by borrowers geographically located in that country and as such include US bank lending to US branches in foreign countries. The data contains information on whether the loans are made to a branch of a foreign bank in that country but does not specify if that branch is a branch of a US bank. UK is the country in our sample which has significant amounts in that category but its presence does not seem to bias the results.
\item[\textsuperscript{15}] The period in our sample is particular in the sense that it begins with the years immediately following the buildup of debt to Latin America and other developing countries and the debt crisis. It would have been interesting to examine the period prior to 1982 because it involved massive lending within short time intervals but our data do not cover this period. See Jain and Gupta (1987) for analysis within this period.
\item[\textsuperscript{16}] Appendix B provides a listing of the large and medium sized banks, including total assets and capital by bank size.
\item[\textsuperscript{17}] We could not locate data for December 1993, so the total number of observations in 40*25=1000.
\end{itemize}
the sample and summary statistics on loan amounts are presented in Table 1. For this sample of
countries, there is a potential source of bias in our estimations as some of the countries listed
experienced a debt crisis leading to a formal rescheduling of debt during the 1982-94 period. During a
debt crisis, protecting outstanding claims often required the involvement of official lenders and the
“concerted lending” of commercial banks.\textsuperscript{19} Debt rescheduling involved the extension of loans that
cannot be viewed as purely voluntary and, respectively, may have been governed by reasons other than
those advanced in this paper. Hence, we provide estimates of the empirical model that exclude
countries with international debt problems as a necessary robustness check.\textsuperscript{20}

3.1 Granger Causality.

In this section we test the hypothesis that large bank lending predicts small bank lending in a panel
of countries and in this sense the results complement those of Jain and Gupta (1987). To test that
proposition, we start by calculating the total amount of outstanding loans (total and short-term) that each
type of bank (large or small) has in each country at 6-month intervals as a proportion of total assets for
the same type of bank. We refer to these as the loan-to-asset ratios for large and small banks. We
then construct a “flow” measure of changes in lending behavior by the two types of banks across
different countries by differencing these data series. The purpose of the exercise is to test whether
lagged “flow” changes in lending by large banks predicts current “flow” changes in lending by small
banks. To carry out any Granger-causality tests, we first investigate the properties of these series. The

\textsuperscript{18} Our sample also excludes former communist countries and offshore banking centers. Turkey and Taiwan are
excluded for lack of sufficient country data.

\textsuperscript{19} See Krugman (1988) for further discussion of private lending to problem debtors.

\textsuperscript{20} Problem debtors were identified from various years of the “World Debt Tables” published by the World Bank.
Twelve of the forty countries in the initial data set were identified as having to reschedule commercial bank debt at
least once during the 1982 to 1994 period. These countries are identified in Table 1 with an asterisk.
time series for each country, however, contain few observations (20) and the results of these tests have to be viewed with caution.

For each country in the sample, we applied Durbin-m to test for serial correlation and to test for unit roots for each country, we performed Augmented Dickey-Fuller tests.\textsuperscript{21} The series appear stationary in all countries (the data series using levels, i.e. loan-to-asset ratios, on the other hand, are non-stationary for most countries in the sample during that period; in other words, levels seem to follow an \textit{I}(1) process) with little evidence of serial correlation.

Ideally, we would have enough observations to run Granger-causality tests for individual countries after transforming the time series in accordance with the characteristics of the series for each country. With pooled data, we use a data set where the data series may exhibit different characteristics in sub-samples. In what follows, we present regression results that allow for both country fixed effects and random effects.

According to the model in section 2, small banks follow large because they act on the information extracted from lagged large bank behavior. Thus large bank lending should predict small bank lending. However, lagged lending by small banks should also predict current large bank lending even if large banks rely on purchased information in each period because the underlying states of the world are correlated across periods. Thus small bank lending may predict large bank lending but, potentially, to a lesser degree.

To test these hypotheses we estimated an equation where current lending (change in loan-to-asset ratios) by each type of bank is predicted by lagged lending by that type of banks as well as by current

\textsuperscript{21} See Hamilton (1994).
and lagged lending by the opposite type of banks.\footnote{We included current lending by the opposite type of bank to capture contemporaneous changes in lending which according to the model is produced by acting on the same information. Also, it is a common practice for large banks to originate syndicated loans with participation by small banks.} We estimated the equation with one, two and three lags and obtained qualitatively equivalent results. In Table 2, we provide the results from estimation with three lags. The results in Table 2 support the hypothesis that there is stronger link between large bank lending and subsequent small bank lending that the reverse. This finding holds for total debt, short-term debt, and short-term debt excluding countries that experienced a debt crisis during the period under examination.

\subsection*{3.2 The effect of persistence.}
The model in section 2 suggests that the extend to which small banks follow large depends, among other things, on the persistence in states over time ($\beta$). To test that proposition, we construct a measure of the persistence parameter $\beta$ as follows. First, we assume that a country’s real GDP per capita (in log form) is directly correlated with the country’s state of world. It follows that a change in that variable from the prior period can indicate a new state of the world, with larger changes being more likely to indicate such a change in the state of the world. To measure these changes, for each country we regress the log of real GDP on its lagged value.\(^{23}\) Residuals between the actual and predicted values become measures of possible changes in the state of the world. To get a measure of the likelihood of a change, we then sum the squared residuals for each regression. The persistence measure for each country is equal to the negative of the residual sums of squares.\(^{24}\)

Table 3 reports the results from regressions in which we tried to capture in a parsimonious form the basic structure of relationships between large and small bank lending. The dependent variables are the net flows of funds (total and for short-term debt only) by small and large banks in the current period and the independent variables are the current and lagged net flows by large and small banks. As stated above, all flows are calculated as the change, within a six month period, in the proportion of assets held in a country by either large or small banks. To test for the effect of persistence $\beta$, we also include the lagged changes in lending by the other type of lender interacted with our persistence measure. For small

---

\(^{23}\) Annual data concerning GDP in real 1990 local units of currency and population were obtained from the IMF *International Financial Statistics* for the period 1970 to 1992. Data before 1970 and after 1992 are not available for some of the countries in the sample.

\(^{24}\) Note that the residual sum of squares proxies for the sample variance of the error term in these regressions. With normally-distributed error terms, an increase in that variance implies that current values contain less predictive power. The formal relation between our measure of persistence and the transitional probabilities $\beta$ and $\varphi$ is presented in Appendix C.
bank lending, the coefficient on this variable should be positive, indicating that small bank lending responds more to lagged changes in lending by large banks when persistence is higher.

The results of testing this hypothesis, reported in Table 3, support such a claim. Namely, a prior increase (decrease) in investment by large banks induced an increase (a decrease) in small bank’s investment in a country in the subsequent period, with the effect being stronger the greater the level of persistence. In contrast, lagged changes in the level of small-bank investments are not linked to subsequent changes in extent of lending by large banks. These results hold for short-term debt for the entire sample and for the sample restricted to countries that did not experience a debt crises during the period in question.

4. Conclusion

This paper has endogenized lenders’ decisions to acquire information in international markets. We have shown that lenders endowed with insufficient wealth, the smaller banks, are those more likely to not purchase additional, country specific information on investment prospects. Instead they will rely on the behavior of more informed investors with a lag to infer information about the international investment prospects. However, the extent of such behavior will depend on the persistence of economic conditions, which varies across countries. Data on international US bank lending support the predictions of the model. Small banks follow the behavior of large banks. Further, such behavior does appear to be more prevalent in countries with more persistence in economic conditions.

At this point, a number of further issues remain to be investigated. First, we have used aggregate economic conditions as a proxy for the likelihood that investment projects will be successful, and thus the persistence in economic conditions relates to the stability of economic conditions in a country. But it is clear that specific borrower information is also important. To address that, we need to identify a
measure for the quality of such borrower-specific information. A second issue is to establish proxies for potential differences in the quality of the signal across countries. The theoretical model makes it clear that such differences will affect the extent of herd behavior.

A third issue involves the distribution of loans between different types of borrower in a country. For instance, the fact that loans to governments may have cross-default clauses such that a default to one borrower triggers repercussions for all creditors can result in different informational content to changes in large firm lending activity to governments as opposed to the private sector. Unfortunately, the data do not offer a breakdown of short-term debt by borrower (government versus private). Some estimates with respect to the link between changes in total (short and long-term) private and government lending by large banks and subsequent changes in small banks’ short term lending suggest that small banks are more responsive to the placement of private rather than government debt by large bank. Also, private lending by small banks appears to respond primarily to private lending by large bank while government lending by small banks does not seem to derive from large bank lending to a particular type of borrowers. Those effects are stronger for countries that exhibit more persistence in economic conditions over time. However, any interpretation of these results should be made with caution because the estimations involve debt of different maturity to different borrowers.
References


<table>
<thead>
<tr>
<th>Country</th>
<th>Average loans of large and small banks for the 1982 to 1994 period (in millions)</th>
<th>Average total loans of large and small banks (Column (1)) divided by their average total assets for the 1982 to 1994 period (in percentages)</th>
<th>Average of small bank loan-to-asset ratio divided by large bank loan-to-asset ratio for the 1982 to 1994 period (in percentages)</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK</td>
<td>25,520</td>
<td>2.01</td>
<td>0.80</td>
</tr>
<tr>
<td>Japan</td>
<td>13,166</td>
<td>1.01</td>
<td>0.48</td>
</tr>
<tr>
<td>Mexico*</td>
<td>6,469</td>
<td>0.51</td>
<td>0.38</td>
</tr>
<tr>
<td>France</td>
<td>6,296</td>
<td>0.51</td>
<td>0.33</td>
</tr>
<tr>
<td>Brazil*</td>
<td>6,111</td>
<td>0.48</td>
<td>0.19</td>
</tr>
<tr>
<td>Belgium</td>
<td>5,554</td>
<td>0.44</td>
<td>0.26</td>
</tr>
<tr>
<td>Canada</td>
<td>5,042</td>
<td>0.40</td>
<td>0.53</td>
</tr>
<tr>
<td>Korea</td>
<td>3,730</td>
<td>0.30</td>
<td>0.31</td>
</tr>
<tr>
<td>Venezuela*</td>
<td>3,582</td>
<td>0.30</td>
<td>0.22</td>
</tr>
<tr>
<td>Italy</td>
<td>3,504</td>
<td>0.28</td>
<td>0.44</td>
</tr>
<tr>
<td>Germany</td>
<td>3,193</td>
<td>0.25</td>
<td>0.22</td>
</tr>
<tr>
<td>Switzerland</td>
<td>3,018</td>
<td>0.23</td>
<td>0.22</td>
</tr>
<tr>
<td>Argentina*</td>
<td>2,939</td>
<td>0.23</td>
<td>0.15</td>
</tr>
<tr>
<td>Sweden</td>
<td>1,949</td>
<td>0.15</td>
<td>0.72</td>
</tr>
<tr>
<td>Netherlands</td>
<td>1,917</td>
<td>0.15</td>
<td>0.30</td>
</tr>
<tr>
<td>Australia</td>
<td>1,839</td>
<td>0.14</td>
<td>0.27</td>
</tr>
<tr>
<td>Spain</td>
<td>1,810</td>
<td>0.14</td>
<td>0.22</td>
</tr>
<tr>
<td>Chile*</td>
<td>1,739</td>
<td>0.14</td>
<td>0.20</td>
</tr>
<tr>
<td>Philippines*</td>
<td>1,732</td>
<td>0.14</td>
<td>0.14</td>
</tr>
<tr>
<td>South Africa</td>
<td>1,532</td>
<td>0.13</td>
<td>0.13</td>
</tr>
<tr>
<td>Indonesia</td>
<td>1,277</td>
<td>0.10</td>
<td>0.07</td>
</tr>
<tr>
<td>Norway</td>
<td>1,046</td>
<td>0.08</td>
<td>0.21</td>
</tr>
<tr>
<td>Denmark</td>
<td>1,026</td>
<td>0.08</td>
<td>0.29</td>
</tr>
<tr>
<td>Colombia*</td>
<td>913</td>
<td>0.07</td>
<td>0.30</td>
</tr>
<tr>
<td>Austria</td>
<td>860</td>
<td>0.07</td>
<td>0.49</td>
</tr>
<tr>
<td>Greece</td>
<td>827</td>
<td>0.06</td>
<td>0.05</td>
</tr>
<tr>
<td>Finland</td>
<td>776</td>
<td>0.06</td>
<td>0.63</td>
</tr>
<tr>
<td>Ecuador*</td>
<td>697</td>
<td>0.06</td>
<td>0.30</td>
</tr>
<tr>
<td>Israel</td>
<td>664</td>
<td>0.06</td>
<td>0.56</td>
</tr>
<tr>
<td>Peru*</td>
<td>575</td>
<td>0.05</td>
<td>0.42</td>
</tr>
<tr>
<td>Ireland</td>
<td>445</td>
<td>0.04</td>
<td>0.19</td>
</tr>
<tr>
<td>India</td>
<td>400</td>
<td>0.03</td>
<td>0.17</td>
</tr>
<tr>
<td>Nigeria*</td>
<td>378</td>
<td>0.03</td>
<td>0.13</td>
</tr>
<tr>
<td>Portugal</td>
<td>345</td>
<td>0.03</td>
<td>0.34</td>
</tr>
<tr>
<td>New Zealand</td>
<td>226</td>
<td>0.02</td>
<td>0.41</td>
</tr>
<tr>
<td>Uruguay*</td>
<td>216</td>
<td>0.02</td>
<td>0.66</td>
</tr>
<tr>
<td>Dominican Republic*</td>
<td>173</td>
<td>0.01</td>
<td>0.40</td>
</tr>
<tr>
<td>Guatemala</td>
<td>94</td>
<td>0.01</td>
<td>3.23</td>
</tr>
<tr>
<td>El Salvador</td>
<td>69</td>
<td>0.01</td>
<td>8.57</td>
</tr>
<tr>
<td>Iceland</td>
<td>69</td>
<td>0.01</td>
<td>0.06</td>
</tr>
</tbody>
</table>

* These 12 countries experienced a consolidation (rescheduling) of commercial bank debt according to World Bank records at least once during the 1984-1992 period.
<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Total debt (long term + short term), all countries</th>
<th>Short term debt, all countries</th>
<th>Short term debt, excluding countries with rescheduled debt agreements</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Small banks</td>
<td>Large banks</td>
<td>Small banks</td>
</tr>
<tr>
<td>Change in large banks’ loan-to-asset ratio for relevant debt type, same period</td>
<td>0.05 (0.02)</td>
<td>0.08 (0.01)</td>
<td>0.07 (0.02)</td>
</tr>
<tr>
<td>Change in large banks’ loan-to-asset ratio for relevant debt type, lag 1 period</td>
<td>0.11 (0.02)</td>
<td>-0.18 (0.06)</td>
<td>0.11 (0.01)</td>
</tr>
<tr>
<td>Change in large banks’ loan-to-asset ratio for relevant debt type, lag 2 periods</td>
<td>-0.11 (0.02)</td>
<td>0.09 (0.04)</td>
<td>-0.08 (0.02)</td>
</tr>
<tr>
<td>Change in large banks’ loan-to-asset ratio for relevant debt type, lag 3 periods</td>
<td>0.01 (0.02)</td>
<td>-0.08 (0.04)</td>
<td>0.02 (0.02)</td>
</tr>
<tr>
<td>Change in small banks’ loan-to-asset ratio for relevant debt type, same period</td>
<td>0.18 (0.07)</td>
<td></td>
<td>0.62 (0.09)</td>
</tr>
<tr>
<td>Change in small banks’ loan-to-asset ratio for relevant debt type, lag 1 period</td>
<td>0.35 (0.03)</td>
<td>0.07 (0.06)</td>
<td>0.34 (0.03)</td>
</tr>
<tr>
<td>Change in small banks’ loan-to-asset ratio for relevant debt type, lag 2 periods</td>
<td>0.25 (0.03)</td>
<td>0.08 (0.06)</td>
<td>0.19 (0.03)</td>
</tr>
<tr>
<td>Change in small banks’ loan-to-asset ratio for relevant debt type, lag 3 periods</td>
<td>-0.19 (0.03)</td>
<td>0.01 (0.05)</td>
<td>-0.21 (0.03)</td>
</tr>
<tr>
<td>Number of observations</td>
<td>760</td>
<td>760</td>
<td>760</td>
</tr>
<tr>
<td>Chi (3)</td>
<td>74.9**</td>
<td>19.6</td>
<td>116.3**</td>
</tr>
</tbody>
</table>

Notes: Standard errors are in parentheses. The Chi (3) statistic is used to test the hypothesis that lagged investment by the opposite type of banks is jointly equal to zero for all lags. **(*) significant at the 0.05(0.1) level. Note that the “relevant” debt type is determined by the debt type of the dependent variable, and is either total debt or short-term debt.
Table 3
Leader--Follower Behavior in International Lending. The Role of Persistence.

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Total debt (long term + short term), all countries</th>
<th>Short term debt, all countries</th>
<th>Short term debt, excluding countries with rescheduled debt agreements</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Small banks</td>
<td>Large banks</td>
<td>Small banks</td>
</tr>
<tr>
<td>Change in relevant loan-to-asset ratio for the opposite type of banks, same period</td>
<td>0.10** (0.02)</td>
<td>0.20* (0.11)</td>
<td>0.09** (1.16)</td>
</tr>
<tr>
<td>Change in relevant loan-to-asset ratio for the opposite type of banks, prior period</td>
<td>0.17** (0.07)</td>
<td>0.03 (0.09)</td>
<td>0.14* (0.01)</td>
</tr>
<tr>
<td>Persistence variable interacted with change for the opposite type of banks, prior period</td>
<td>71.2** (33.1)</td>
<td>76.1 (85.8)</td>
<td>55.8** (13.3)</td>
</tr>
<tr>
<td>Number of observations</td>
<td>840</td>
<td>840</td>
<td>840</td>
</tr>
<tr>
<td>Model Chi (3)</td>
<td>23.78</td>
<td>28.46</td>
<td>727.6</td>
</tr>
</tbody>
</table>

Notes: Robust estimation. In addition, the estimation procedure incorporates fixed and random effects. Standard errors in parentheses. The terms ** (*) indicate significance at the 0.05 (0.1) level. When flow variables are constructed, the first, the last and two observations around December 1993 are lost so the sample becomes 40 times (25 - 4), or 840. Recall that we were not able to obtain data for December 1993. Note that the “relevant” loan to asset ratio is calculated using the loan type of the dependent variable, either total loans or short-term loans. Excluded countries with debt rescheduling agreements are identified in Table 1.
Appendix A

Proof of Propositions 1 and 2

Consider two processes: \( q_L \) and \( q_H \) such that \( q_H > q_L \). We want to establish that \( t^*/T \) is smaller when \( q = q_H \). We have established that the slope of the path of \( \alpha \) increases in and that greater \( q \) implies a greater amplitude of the dynamic path. Consequently, in expected terms, the two paths have a single intersection point. Denote that value by \( \alpha = \alpha^* \). Let \( \alpha_D < \alpha_L < \alpha_H < \alpha_U \) where \( \alpha_U \) and \( \alpha_D \) are the upper and lower bounds of the process \( q=q_L \). Let \( t_{HH} \) be the time when \( \alpha = \alpha_H \) under \( q = q_H \), let \( t_{LL} \) be the time when \( \alpha = \alpha_L \) under \( q = q_L \), let \( t_{HL} \) be the time when \( \alpha = \alpha_H \) under \( q = q_L \) and let \( t_{LH} \) be the time when \( \alpha = \alpha_L \) under \( q = q_H \).

**Proposition 1:** Let \( S=H \). We need to consider three cases:

- **case 1:** \( \alpha_D < \alpha_L < \alpha_H < \alpha_U \). In this case, \( t_{LH} - t_{LL} > 0 \) and \( t_{HH} - t_{HL} > 0 \).
- **case 2:** \( \alpha_D < \alpha_L < \alpha^* < \alpha_H < \alpha_U \). In this case, also, \( t_{LH} - t_{LL} > 0 \) and \( t_{HH} - t_{HL} > 0 \).
- **case 3:** \( \alpha_D < \alpha^* < \alpha_L < \alpha_H < \alpha_U \). In this case, also, \( t_{LH} - t_{LL} < 0 \) and \( t_{HH} - t_{HL} > 0 \).

However, by concavity \( |t_{LH} - t_{LL}| < |t_{HH} - t_{HL}| \). Hence, in all three cases \( t^*/T \) is greater under \( q_H \).

By symmetry, the above argument applies for \( S=L \), and, thus for the overall dynamic paths.

**Proposition 2:** Consider two processes with equal \( q \) and different persistence \( \beta_H \) and \( \beta_L \) such that \( \beta_H > \beta_L \). Under the first process, each state is expected to persist for \( \beta_H / (1 - \beta_H) = t_H \) and under the second process, for \( \beta_L / (1 - \beta_L) = t_L \) so that \( t_H > t_L \). Take a time period \( T^* = \text{lcm}(t_H, t_L) \), i.e. in \( T^* \) periods, the first process will switch \( T^*/t_H \) times and the second process will switch \( T^*/t_L \) times, and at the end of the \( (t + T^*) \)th period, the two processes will be in the same starting position as at time \( t \).

Now, the duration of each state under the high \( \beta \), \( t_H \), can be separated into two parts: \( t_{H1} = t_L \) and \( t_{H2} = t_H - t_L \). In other words, \( t_{H1} \) are times after the two processes have switched simultaneously until the
time when the process with low $\beta$ switches again; $t_{H2}$ are the times when the low $\beta$ has switched but the high $\beta$ process has not.

In the case of $t_{H2}$ times, it is obvious that $\alpha$ will have unambiguously greater/smaller values under $\beta_H$ as compared to $\beta_L$ when the state is H/L. We need to show that $t^*/t_{H1}$ is greater under the $\beta_H$ process. As with Proposition 1, we proceed by considering different cases, and the same reasoning applies. However, here, we have an additional case, one in which the two dynamic paths do not have an intersection point within the $t_{H1}$ periods. That we denote as case 4: After a simultaneous switch, the paths of $\alpha$ under $\beta_L$ and $\alpha$ under $\beta_H$ do not have an intersection point. Then, $t_{LH} - t_{LL} > 0$ and $t_{HH} - t_{HL} > 0$. This completes the proof.

In a word, in $T^*$ periods, each switch of the $\beta_H$ process is matched by a switch of the $\beta_L$ process. It takes $t_{H1}$ periods until the $\beta_L$ process switches again. We have shown that within these $t_{H1}$ periods, $\alpha$ “spends more time” in the region outside $(\alpha_L, \alpha_H)$. For the remainder of the time, i.e. the $t_{H2}$ times, the process under $\beta_H$ is in the tails region, while the process $\beta_H$ switches from one state to another.
Appendix B

Identifying “Large”, “Medium-Sized” and “Small” Banks
Bank Names, Assets and Capital as of June 31, 1992

<table>
<thead>
<tr>
<th>Large Banks</th>
<th>Medium-Sized Banks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bank of America</td>
<td>Wells Fargo</td>
</tr>
<tr>
<td>Citibank</td>
<td>Marine Midland</td>
</tr>
<tr>
<td>Chase Manhattan Bank</td>
<td>Mellon Bank</td>
</tr>
<tr>
<td>Morgan Guaranty</td>
<td>First National of Boston</td>
</tr>
<tr>
<td>Chemical Bank</td>
<td>National Bank of Detroit</td>
</tr>
<tr>
<td>Continental Illinois</td>
<td>Texas Commerce Bank</td>
</tr>
<tr>
<td>Bankers Trust</td>
<td>Bank of New York</td>
</tr>
<tr>
<td>First National of Chicago</td>
<td>Nationsbank Texas</td>
</tr>
<tr>
<td></td>
<td>Republic National of New York</td>
</tr>
<tr>
<td></td>
<td>First Interstate of California</td>
</tr>
<tr>
<td></td>
<td>First City National of Houston</td>
</tr>
</tbody>
</table>

The names of small banks are not provided.

Total Assets: Large Banks: $694.1 billion
              Medium Banks: $273.0 billion
              Small Banks: $766.0 billion

Total Capital: Large Banks: $68.0 billion
                Medium Banks: $26.4 billion
                Small Banks: $71.3 billion
Appendix C

The Formal Link Between the Persistence Parameter $\beta$ and Root Mean Square Errors

Below we demonstrate that a higher root mean squared error implies less persistence in economic conditions as reflected in the model by a lower $\beta$. In particular, assume that $Y_t = a + b Y_{t-1} + e_t$ where $e_t$ is distributed $n(0, \sigma^2)$. Substituting into our expression for persistence, $\beta$, we obtain:

\[
\beta = \text{Prob}(Y_t \geq Y^* | Y_{t-1} \geq Y^*) \\
= \text{Prob}(a + b Y_{t-1} + e_t \geq Y^* | Y_{t-1} \geq Y^*) \\
= \text{Prob}(e_t \geq Y^* - a - b Y_{t-1} | Y_{t-1} \geq Y^*) \\
= 1 - \text{Prob}(e_t \geq Z_t | Y_{t-1} \geq Y^*) 	ext{ where } Z_t = Y^* - a - b Y_{t-1} \leq 0.
\]

Thus,

\[
\beta = 1 - \int_{Y^*}^{\infty} \int_{Y^*}^{\infty} dG(e) dF(Y)
\]

(2)  $\beta = 1 - \int_{Y^*}^{\infty} \int_{Y^*}^{\infty} dF(Y)

In the empirical analysis, we use as our empirical proxy for $\beta$ the inverse of the root mean squared errors ($1/\text{RMSE}$) for each country. Note that:

\[
\text{RMSE} = \frac{\sum_{t=1}^{T} (Y_t - \bar{Y}_t)^2}{T} = \sigma^2
\]

Below we demonstrate that, for any $Z_t > 0$ and any two normal distributions $G_1(e)$ and $G_2(e)$ with means zero and variances $\sigma_1^2$ and $\sigma_2^2$ such that $\sigma_1^2 > \sigma_2^2$, and thus, $\text{RMSE}_1 > \text{RMSE}_2$, we have that

\[
\int_{Z_t}^{\infty} dG_1(e) > \int_{Z_t}^{\infty} dG_2(e)
\]

and thus $\beta_1 < \beta_2$. Let $g_1(e)$ and $g_2(e)$ be the corresponding density functions. Define $e^*$ such that $g_1(e^*) = g_2(e^*)$:

\[
e^* = \pm \sqrt{\frac{2\sigma_1^2 \sigma_2^2}{\ln \left( \frac{\sigma_1}{\sigma_2} \right) \left( \sigma_1^2 - \sigma_2^2 \right)}}
\]

For a standard normal distribution:
(6) \[ g_1(0) = \frac{1}{\sigma_1 \sqrt{2\pi}} < \frac{1}{\sigma_2 \sqrt{2\pi}} = g_2(0) \]

From (5), (6) and \( \frac{dg_i(e)}{de} < 0 \), \( i = 1,2 \) we have that

(7) \[ g_1(e) > g_2(e) \quad e > e^* \]
\[ g_1(e) \leq g_2(e) \quad e \leq e^* \]

Since \[ \int_0^\infty g_1(e)de = \int_0^\infty g_2(e)de = \frac{1}{2} \], it follows that

(8) \[ \int_0^{e^*} g_2(e)de - \int_0^{e^*} g_1(e)de = \int_0^{e^*} g_1(e)de - \int_0^{e^*} g_2(e)de > 0 \]

As indicated below, (4) holds for the case of \( 0 < Z_i < e^* \):

(9) \[ \int_{Z_i}^{\infty} g_1(e)de - \int_{Z_i}^{\infty} g_2(e)de = \int_{Z_i}^{e^*} \left[ g_1(e) - g_2(e) \right] de - \int_{Z_i}^{e^*} \left[ g_2(e) - g_1(e) \right] de > 0 \]

Similarly, for \( Z_i \geq e^* \), we have:

(10) \[ \int_{Z_i}^{\infty} g_1(e)de - \int_{Z_i}^{\infty} g_2(e)de = \int_{Z_i}^{\infty} \left[ g_1(e) - g_2(e) \right] de > 0 \]

Thus, (4) holds, and \( \beta \) will be inversely related to the RMSE, and directly related to the inverse of the root mean squared error. Note that the above result depends on the normality assumption. If the distribution is not normal, (7) may be violated.