A model of bank contagion through lending

Patrick Honohan a,b,*

a Development Research Group, The World Bank, 1818 H Street NW, Washington, DC 20433, USA
b Centre for Economic Policy Research, London, UK

Received 1 October 1997; accepted 22 May 1998

Abstract

While the theoretical literature on bank contagion has focused mainly on deposit withdrawals, a disturbance on the lending side can also propagate through the system. We find that commonly observed features of banking crises can be rationalized in terms of a simple oligopoly model of bank lending. Informational externalities can result in banks over-committing to risky lending, and finding themselves in difficulties even if they were not consciously taking advantage of the deposit-put option implicitly offered by government guarantees. Liberalization of entry can also disturb an oligopolistic equilibrium and induce banks to shift to a riskier stance—this effect will be stronger if the old banks have a higher cost-base with quasi-fixed factors of production. © 1999 Elsevier Science Inc. All rights reserved.

Keywords: Banking crisis; Information externalities; Contagion

1. Introduction

When systemic banking crises strike, they may have been caused by system-wide shocks, or by a common behavioral response to perverse incentives. But the possibility that a problem in one segment of the system, or even in one institution, may propagate and infect the system more widely must also be considered.

Observers of recent banking crises have documented the herding behavior of banks with associated euphoria and over-lending, which has often preceded the collapse. They have also noted the scramble for return at the cost of higher risk, as existing banks struggle to stay in business in the face of new competition in a liberalized financial market. Both of these important phenomena reflect a type of contagion that has been overshadowed by the theoretical literature that focuses on depositor runs.

Indeed, the banking crises of the past quarter-century in industrial countries have
been characterized more by correlated failures in credit quality than by depositor runs. The purpose of this paper is to describe how disturbances in one part of the banking system are transmitted through lending decisions. To this end, we develop a simple model of oligopolistic banking and describe the equilibrium risk of the banks’ portfolio. Small disturbances directly affecting only one bank can propagate through the system. We show in particular how the arrival of a new entrant can lead to riskier behavior by all banks, and how false information leading to over-optimism by one bank can lead to over-optimism by all.

An extensive literature has considered the structure and stability of banking systems. Banking forms a network in which a relatively small number of institutions interact to provide credit and process information (Shubik, 1990). Disturbances to part of such a network can have far-reaching and hard-to-predict consequences for the whole (Honohan & Vittas, 1997). While financial networks appear to be robust to most disturbances (Benston & Kaufman, 1995), positive feedback could cause catastrophic collapse in some circumstances. The fear of contagious propagation of disturbances through the financial system has influenced policy structures such as (on the one hand) government deposit guarantees and (on the other) the imposition and supervision of restrictive regulations on bank conduct. The degree to which contagion actually presents a serious risk in practice is debated (Kaufman, 1994). While the theoretical literature on bank contagion has focused mainly on deposit withdrawals as a propagation mechanism (cf. Diamond & Dybvig, 1983; Postlewaite & Vives, 1987; Chari & Jagannathan, 1988; and Calomiris & Kahn; 1991), the empirical and historical literature also mentions that a disturbance on the lending side can propagate through the system (cf. Davis, 1992; Honohan, 1997).

To capture some of the features of how competitive pressures on bank decision-making and risk-taking propagate, this paper considers a simple oligopoly model of bank lending. For heuristic purposes, we begin in Section 2 with an intuitive account of the model and the results obtained. We show that both the commonly observed features of herding behavior and the acceptance of increased risk as a response to increased competition can be rationalized in terms of such a model.

In the basic model, presented in Section 3, each of a small number of banks faces a downward sloping demand curve from low-risk borrowers. The slope of this curve is common knowledge and banks compete with Cournot strategies to determine their volume of lending. Banks also have an alternative high-risk lending category. Cournot competition occurs in the market for high-risk loans, too. We assume that the supply of low- and high-risk loans are independent, and that banks can identify which borrowers are high-risk. Banks also incur costs, not only the interest costs of deposits, but resource costs in administering, processing and monitoring. These costs will vary with the size of the balance sheet. Finally, the supply of deposits to the bank is not dependent on the risk of failure (implicitly because the deposits are regarded as being guaranteed by the unmodelled government).

We adapt this set-up to consider the effect of two distinct types of disturbance on system-wide exposure or risk of failure.

In Section 4, we acknowledge that banks’ information about the degree of risk
involved in the high-risk loans is imperfect, and that each bank attempts to infer the degree of risk from observation of market behavior. We show that plausible learning processes can amplify the risk of systemic failure, relative to a situation where no learning takes place. Informational externalities can result in banks over-committing to high-risk lending, and finding themselves in difficulties—even if they were not consciously taking advantage of the deposit-put option implicitly offered by government guarantees.

In Section 5, we consider the role of liberalization, modeled as new entry, in circumstances where existing banks are locked into a certain level of resource costs; in other words, there are quasi-fixed factors of production. We show how entry, disturbing the previous oligopolistic equilibrium, will tend to increase the amount of risk borne by the system above what would occur in equilibrium. This effect will be stronger if the old banks have a higher cost base. The model suggests that the impact effect of liberalization in this regard is likely to be more severe than its long-term effect.

2. An Intuitive Account

Bank lending typically operates in an environment that is not perfectly competitive; banks deal in assets that are not priced on an open market. These are two key features that we try to capture in the modeling approach adopted in this paper. They allow us to see how changes in the environment can swing a banking system from a position of safety to one in which failure becomes likely.

In particular, the fact that the true value of assets is revealed slowly allows misinformation to propagate uncorrected in the market, thereby increasing the risk of failure.

Many recent studies have focused on the deposit-put: the fact that a bank offering to accept deposits that are—explicitly or implicitly—insured by the government, is only promising to pay back if it remains solvent. By manipulating the risk of failure, the bank can effectively reduce the total expected cost of servicing these deposits. In other words, through the deposit insurance, the bank has acquired a put option on the value of the deposits, an option whose value increases with the risk of bank failure.

The present model also embodies a deposit put option, but it is one whose value may, in certain conditions, be zero. Changes in the competitive environment, for example new entry by a low-cost bank, may shift the existing banks from a position where failure is impossible (and the deposit-put worthless) to one where failure is a real possibility; the banks, therefore, may begin to exploit the deposit-put in a socially costly way.

Each of the several banks in our model is seen as having unlimited access to deposits paying an exogenously fixed interest rate. (That the deposit interest rate does not depend on the risk of failure reflects the implicit assumption that deposits are being guaranteed by the unmodeled government). The banks must hold a minimum capital ratio, and they can access capital at an exogenously fixed cost, higher than that on deposits. They can invest their resources in loans, distinguishing between low- and high-risk loans. We allow for a constant per-dollar unit administrative cost of lending, to capture information processing, screening and monitoring costs. The low-risk loans
yield a known return which declines with the size of aggregate lending. The elasticity
of this low-risk loan supply curve is common knowledge.

The high-risk loans pay an exogenously fixed return if successful, but the probability
of success declines with the volume of aggregate lending. We make various assumptions
about participants’ information about the elasticity of this probability of success curve.
This declining probability of success is a short-hand intended to reflect an underlying
reality which we can picture as beginning with the banks being faced with a pool of
applicants. Some of the applicants are identifiable as low-risk. Efforts are made to
screen the remainder, but (presumably because the most plausible applicants are
granted loans first) the success rate of screening deteriorates as the aggregate number
of loans to this category increases. We assume that the supplies of low- and high-risk
loans are independent, and (as mentioned) that banks can identify which borrowers
are high-risk.

We model the market equilibrium as a Cournot-type, where each bank optimizes
conditional on the lending decisions of the others. There are thus two important
downward sloping market demand curves: that coming from low-risk borrowers and
that from the high-risk borrowers. The two elasticities of these curves are determinants,
along with the exogenous interest rates on deposits and capital, of the equilibrium of
the system and, in particular, of the equilibrium ratio of low- to high-risk loans.

If the capital adequacy requirements are sufficiently high, we are in the no-failure
zone and no bank will fail even if all of its high-risk borrowers default. If capital
requirements are lower, then banks will recognize that there is a possibility of failure.
Recognizing that this gives them the possibility of exercising the deposit-put, they
will increase the share of high-risk loans, thereby increasing the probability of failure.
The precise conditions under which this occurs are spelled out in the next section.
So far, nothing has been said about contagion as such (in particular, a bank can be
in the failure zone with no contagion existing).

This basic framework is then used to explore the consequences of informational
disturbances and entry. Up to now, we have assumed that the banks know the two
market demand curves. Each bank may have some imperfect information, and this
will provide an incentive for all to observe each others’ behavior in order to try to
infer what the others know. This kind of problem is widely discussed in connection
with the determination of market prices of assets when information is imperfect and
unevenly distributed. In that case, a key question is the degree to which all of the
information in the market becomes revealed in prices. But here the assets are not
priced, so that information does not feed back directly onto market prices.

We extend our model to take into account imperfect information, and derive the
new market equilibrium conditions. We show, under a simple but plausible learning
mechanism, how false information received by one bank can result in erroneous
lending decisions throughout the system. Essentially, we show that each bank can
infer the others’ beliefs from their behavior. If each adjusts its own beliefs in a quasi-
Bayesian fashion based on what it observes about the beliefs of others, each bank’s
beliefs can converge to a common value. Because there is no feedback onto market
prices, the erroneous lending pattern can persist for some time. For example, excessive
optimism by one bank can result in too much risky lending throughout the system. That certainly rings true for the kinds of bandwagon effect or herd behavior observed in credit-sustained property booms.

The effects of entry can also be modeled as a disturbance to the Cournot equilibrium. Here, the interesting point is that the arrival of a new bank can move the equilibrium out of the no-failure zone, leading not only to a risk of failure, but to the exploitation of the deposit-put and increasingly risky behavior. The problem is that, having inherited staff levels that justify a certain scale of lending, incumbents will treat these as sunk costs and continue to lend at a level which is beyond what is now prudent, given the lower returns and higher risk that is emerging in the market as a result of the arrival of the new entrant.

In the long-run, when the fixed costs incurred by the old banks are no longer relevant, it is possible that the system could return to the no-failure zone. To that extent, the disturbing effect of entry may vanish over time, and this disturbance may well be an acceptable cost to pay for a more efficient long-term equilibrium.

3. The Basic Model

To begin, we describe those features of the set-up that are common to both variants of the model. There are \( n \) banks in the market. Each bank \( i \) can access deposits at an exogenously fixed return \( r^0 \) (interest rate \( r^0 - 1 \)). It can supply loans at a per-dollar unit administrative cost \( w_i \). There are two types of loan: low- and high-risk. The low-risk loans (risk-free loans) yield a certain return \( r_L \). Each bank’s portfolio of high risk loans yields \( r_H \) if repaid, nothing otherwise. The probability of repayment is denoted \( \theta \). We assume this risk is non-diversifiable; the entire risky portfolio of a given bank either repays or does not. The opportunity cost of capital is exogenously fixed at \( r_C \). It is assumed that the maximum yield on a high-risk loan \( r_H \) is exogenously fixed, but the low interest yield \( r_L \) and the probability of repayment \( \theta \) depend on the aggregate volume of loans of each type. We write \( x_i \) and \( z_i \), \( X \) and \( Z \), as the quantity of low-risk and high risk loans of bank \( i \) and the aggregate of the banks respectively. Thus the dependence of the returns on volume is: \( r_L = f(X) \) and \( \theta = g(Z) \) with elasticities: \( \varepsilon = Xf'(X)/r_L \) and \( \eta = Zg'(Z)/\theta \).

Banks are required to hold a minimum capital-to-assets ratio of \( \gamma \), and we will assume that the ex ante franchise value of banking is sufficient to remunerate this capital, but that the opportunity cost of capital \( r_C \) is higher than \( r^0 \) so that the minimum capital ratio is binding. The weighted average cost of funds is thus [Eq. (1)]:

\[
    r^* = (1 - \gamma) r^0 + \gamma r_C. \tag{1}
\]

3.1. Equilibrium in the no-failure zone

If the capital requirement is set sufficiently high so that the probability of failure is zero in the relevant range—in other words if we are in the “no-failure zone”—then the appropriate maximand is:

\[
    \pi_i = \theta r_H z_i + r_L x_i - (r^* + w_i)(z_i + x_i) \tag{2}
\]
Maximization of the expected value of Eq. (2) by each bank under the Cournot assumption that the volume decisions of the others are not sensitive to one’s own, leads to the familiar Lerner price-cost ratio conditions for each bank in the two markets:

\[
1 - \frac{r^* + w_i}{r^L} = -\frac{x_i}{X} \epsilon
\]

(3)

\[
1 - \frac{r^* + w_j}{\theta r^H} = -\frac{z_i}{Z} \eta
\]

(4)

The intuition here is that each bank takes into account the impact of its own lending on the rate of return; thus it will expand its lending only to the point where, conditional on the lending of all other banks, a further increase would lower the gross margin on other loans by more than it would generate in new net revenue. These conditions Eqs. (3) and (4) can thus be seen as reaction functions\(^9\) determining each bank’s loan supply \(x_i\) and \(z_i\) as a function of the lending by others: \((X - x_i)\) and \((Z - z_i)\). Aggregating these over all of the banks provides the market equilibrium conditions:

\[
1 - \frac{r^*X + \sum w_i x_i}{r^L X} = -\sum \frac{x_i^2}{X^2} \epsilon
\]

(5)

and

\[
1 - \frac{r^*Z + \sum w_i z_i}{\theta r^H Z} = -\sum \frac{z_i^2}{Z^2} \eta
\]

(6)

The market equilibrium conditions Eqs. (5) and (6) thus determine the risk-free (low-risk) lending rate \(r^L\) and the risk parameter \(\theta\). In the case where all banks are identical, the sum of squared market shares (Herfindahl index) is simply the reciprocal of the number of banks, so the market equilibrium conditions reduce to [Eqs. (7) and (8)]:

\[
1 - \frac{r^* + w}{r^L} = -\frac{\epsilon}{N}
\]

(7)

\[
1 - \frac{r^* + w}{\theta r^H} = -\frac{\eta}{N}
\]

(8)

In words, oligopolistic competition between the banks will drive loan supply up to the point where the expected profit margin, as a percentage of gross return on each sub-portfolio, is inversely proportional to the price elasticity of the loan supply. These conditions are, of course, directly analogous to classic IO results defining the profit share as directly proportional to the absolute value of the elasticity of demand and inversely proportional to the number of firms.

A small number of banks and a high capital requirement means that the gross return on low-risk assets is high: writing the interest and administrative costs devoted per dollar of lending as:

\[
b_i = (1 - \gamma) r^0 + w_i
\]
we may regard the gross profit per dollar lent \( r^L - b_i \) as a measure of the franchise value of low-risk banking.

3.2. Equilibrium in the failure zone

If the capital adequacy requirements are not sufficient to rule out failure, then the optimization criteria change. This is because, by increasing the share of high-risk loans in its portfolio, each bank can influence the value of the deposit-put, the possibility that losses by the bank in excess of its capital (thereby threatening the depositors) will be met by the government.

If the high-risk loan is not repaid, then the government may have to pay out to depositors. The government is at risk of having to make such a pay-out if the sum due to the depositors and factors of production (the latter assumed already paid out of depositor funds) exceeds the amount received by the bank from the low risk borrowers, i.e. the government is at risk if Condition F holds, defined by the inequality

\[
b_i(x_i + z_i) - r^L x_i \equiv F_i > 0.
\]

As the deposit-put will only have value if this inequality is satisfied, we say that a bank is ex ante in the failure zone whenever Condition F holds. \( F_i \) is effectively the value of the deposit put if exercised. If Condition F does not hold, then the bank will still be able to pay depositors even if the high-risk loan is not repaid; it is not in the failure zone, and the deposit put has no value.

Note that Condition F will not hold if what we have termed above the franchise value of low-risk banking \( r^L - b_i \) is sufficiently high.

The bank maximizing expected profits in the failure zone will thus maximize [Eq. (9)]

\[
\pi_i + (I - \theta)F_i = \theta (r^H z_i + r^L x_i) - (r^* + w_i)(x_i + z_i)
\]

\[
+ (I - \theta)(b_i(x_i + z_i) - r^L x_i)
\]

\[
= \theta (r^H z_i + r^L x_i - (b_i + \gamma) (x_i + z_i)).
\]

\[
= \theta (r^H z_i + r^L x_i - \theta(r^* + w_i + a)(x_i + z_i),
\]

where,

\[
a = \gamma r^C (I - \theta)/\theta.
\]

In examining this expression it is worth noting that, in a failure zone, the shareholders of a failing bank can lose their entire capital, including the profit from the low-risk loans. On the other hand, from the bank’s (shareholder’s) viewpoint the costs, other than opportunity cost of capital, also become contingent on not failing. Although the precise formulation is an implication of the very simple dichotomous risk specification, the insight that the risk differential between the two loan categories is reduced seems robust. An increase in \( \theta \) therefore affects the expected return not only on the high-risk asset, but also on the low-risk asset and the effective cost of lending.

With this maximand and Cournot behaviour, conditions Eqs. (3) and (4) above for individual bank asset choices in the no-failure zone are replaced by failure zone conditions Eqs. (10) and (11):
\[
1 - \frac{r^* + w_i + a}{r^L} = -\frac{x_i}{X} \epsilon \tag{10}
\]

and
\[
1 - B_i \frac{r^* + w_i}{\theta r^H} = \frac{z_i^*}{Z}, \tag{11}
\]

where
\[
B_i = \frac{1}{\left(1 + \frac{b_i}{\gamma r^L}\right)} \left(1 - \frac{b_i}{r^H}\right).
\]

modifies the mark-up ratio, and is a grossed-up risk asset quantity, effectively taking account of the increased risk of failure, and
\[
z_i^* = z_i + \frac{r^L - b_i}{r^H - b_i} x_i
\]

hence of losing the profits from the risk-free asset, imparted by additional purchases of the high-risk asset.

The only difference between Eq. (3) and Eq. 1(0) is that the cost term in Eq. (10), i.e. \( r^* + w_i + a \), is evidently larger than that in Eq. (3). Thus the existence of the deposit-put acts just like an increase in the cost of capital.

The relation between Eq. (4) and Eq. (11) is more complex. The factor \( B_i \) by which the LHS of Eq. (11) differs from that of Eq. (4), tends to zero for small capital requirements (\( \gamma \to 0 \)) making the LHS of Eq. (11) larger than that of Eq. (4). Accordingly, the grossed-up risk assets total is higher than the total satisfied:
\[
z_i < z_i^* < z_i + x_i.
\]

Comparing the incentives of an individual bank in the failure zone with that of the same bank not in the failure zone, conditional on market prices \( r^L \) and \( \theta \), we see that the failure zone bank will place slightly less in low-risk assets. Furthermore, its grossed-up risk position, \( z_i^* \), will be higher provided that \( B_i < 1 \); and the net amount placed in high-risk assets, \( z_i \), will also be larger, unless what we have termed the franchise value of low-risk banking is sufficiently high. The condition\(^{10} \) under which the bank will place more in high-risk assets will be referred to as Condition \( R \).

The market equilibrium condition for identical banks in the failure zone is a straightforward aggregation of Eq. (10) and Eq. (11). Therefore, for the same elasticity \( \epsilon \), the equilibrium low-risk rate \( r^L \) is unambiguously higher, though by a small amount if capital requirements \( \gamma \) are small. Equilibrium in the high-risk asset market will involve (if elasticity \( \eta \) is little changed and under condition \( R \)) increased lending and lower probability of success \( \theta \).

One way of summarizing these findings is to say that, comparing the equilibrium of a banking system with unlimited liability with a system such as we have described in the failure zone, there is:
i. an increase in the low-risk lending rate sufficient to offset the relatively higher burden being borne by the capital committed to “low-risk” lending, but which will be lost if the bank fails;

ii. an expansion of high risk loans, to the point where the value of $\theta$ has fallen sufficiently to ensure that $\theta r^H$ gives the same margin over the lower effective marginal cost of lending.

Note that we have not yet discussed the question of contagion, to which we now turn.

4. Informational Disturbances

In this section, we assume that the banks do not know $\eta$, but make inferences about it from (a) observing a noisy private signal and/or (b) observing the behavior of other banks in response to their noisy private signal. Because the market price of high-risk loans does not adjust (the probability of loss adjusts, but this is not observed) the learning process is rather different to that analyzed in models of learning and price revelation in securities markets.

We analyze a Cournot equilibrium of quantities and expectations, in which all banks share common information. In order to obtain closed solutions in a stochastic environment, we use a log-linear functional form for the dependence $g$ of loss probability on the volume of lending. Specifically, we assume that [Eq. (12)]:

$$
\theta = \exp\{-\beta Z\}
$$

which implies that the absolute value of elasticity $\eta$ increases with total lending $Z$:

$$
\eta = -\beta Z.
$$

The parameter $\beta$ is not known, but that banks have probabilistic beliefs about it represented by a Normal (Gaussian) distribution with mean $\zeta$ and variance $\sigma^2$.

We begin by examining the equilibrium behaviour conditional on the banks’ beliefs; we will return below to the determination of these beliefs. Thus, given $\zeta$ and $\sigma$ [Eq. (13)]:

$$
\theta = (2 + \sigma^2 Z^2/2) \exp\{-\zeta Z\}.
$$

Maximizing the expected value of $\pi^*$ using this expression for the expected value of $\theta$ leads to the choice of $z_i$ being determined by:

$$
1 - \frac{r^* + w_i}{\theta r^H} = \frac{z_i}{Z} \eta^*
$$

where [Eq. (15)],

$$
\eta^* = \frac{Z}{\theta} \frac{\delta \theta}{\delta Z} = -\zeta Z + \nu
$$

and [Eq. (16)],

$$
\nu = (\zeta Z + 2) \frac{\sigma^2 Z^2/2}{1 + \sigma^2 Z^2/2}
$$
In words, the equilibrium mark-up condition has the same general form as in the certainty case, but with the expected loss rate replacing the known loss rate, and with a “risk-adjusted elasticity factor” replacing the known elasticity. The risk-adjusted elasticity factor differs from the expected value of the elasticity $\xi Z$ by a term roughly proportional to the variance $\sigma^2$ of the estimate of the parameter $\beta$.

Note that the individual bank condition Eq. (14) implicitly defines an $z_i$ as an invertible function of $z$, since the left hand side is a decreasing, and the right hand side an increasing, function of $\xi$.

In the failure zone, where loss will result in failure and the deposit-put, we substitute an expression like Eq. (11) for Eq. (14) again modified by replacing $\theta$ with $\bar{\theta}$ and $\eta$ with $\eta^*$.

The market equilibrium condition is immediately [Eq. (17)]:

$$1 - \frac{r^* + \sum w_i}{\bar{\theta} + \eta^*} = - \sum \frac{z_i}{\xi Z}$$

It follows that an increase in uncertainty $\sigma$ reduces the absolute value of the risk-adjusted elasticity term so that, for the same expected value of the parameter $\beta$, the mark-up will be smaller. (This is a small effect which depends on the non-linear manner in which the parameter $\beta$ enters the decision rule; it is dependent on the chosen functional form for $g$ and might not be very robust to alternative functional forms).

We turn now to the formation of beliefs concerning $\beta$. Suppose that all the banks share the same subjective variance $\sigma$ of the parameter $\beta$, and that this is constant. Since this is a one-period model, there is no chance for banks to update their beliefs on the basis of loan outcomes. But if the banks can observe the choice $z_i$ and the cost level $w_j$ of each of the other banks $j$ then each bank is able to infer the value of each other bank’s $\xi$ by solving the individual optimization formula Eq. (14).

For example, if it is common knowledge that each bank observes a private signal $\zeta_i$ with the property that:

$$\zeta_i = \beta + u_i$$

where the $u_i$ are independent normally distributed with zero mean and a common variance $\sigma^2$, then if each bank can observe what the others would lend on the basis of their private information, then each bank can learn about the private signals and each can form the improved estimate $\zeta = \sum \zeta_i / n$ of $\beta$ with variance $\sigma^2$. Accordingly, a configuration in which all banks do not optimally use this improved estimate cannot be an equilibrium. Specifically, each bank $i$ will combine its own information $\zeta_i$ with the observed $\sum_{j \neq i} \zeta_j / n$ for the others, a process consistent with equilibrium at a common value $\zeta = \sum \zeta_i / n$.

The restrictive nature of the Cournot assumption is relevant here in two respects. First, it implies no strategic signaling behavior. Second, it begs the question of dynamics: would a process of learning each other’s signals converge to full revelation?

This second point can be addressed fairly readily with the following simple example. Suppose it takes some time (several “weeks”) for the lending process to be accomplished, but that each bank’s intended lending, and hence its own parameter $\zeta_i$, can
be observed. Suppose that each bank $i$ updates its parameter value $\zeta_i$ weekly, beginning with its private signal $\zeta_i^0$, and then (as it observes the parameters $\zeta_j$ of others) updating by linearly combining the last period’s parameter $\zeta_j^{t-1}$ with this week’s mean of other banks’ revealed parameters, and then iterating weekly on this updating procedure. That is,

$$\zeta_i^t = \gamma \frac{1}{n} \sum \zeta_j^{t-1} + (1 - \gamma)\zeta_i^{t-1}$$

It can readily be shown that this process will converge to the mean of the private signals with probability one, as the eigenroots of the associated transition matrix have $n - 1$ roots equal to $1 - \gamma$, and one root at unity.

In contrast to models of price signal aggregation like that in Diamond and Verrechia (1981), here the increased loan supply does not reflect back onto an observed market price until after the lending has occurred. Learning is not corrected by a global market feedback. That implies that false signals can result in banks unwittingly exposing themselves to a high risk of failure. In this case, therefore, failure can occur even if banks have not employed the deposit-put modification of Eq. (9). Furthermore, this signal propagation, along with the oligopolistic market equilibrium behavior, implies that a false positive signal received by one bank will result in all banks adopting a riskier strategy and an increase in the failure probability of banks other than the one which received the signal.

It is worth stressing again that the Cournot assumption means that each bank neglects the impact of its behavior on the others. Insofar as an expansion of risky lending in fact induces more risky behavior all around, alternative strategic assumptions, perhaps damping the system’s response to extreme drawings of the signal, could be explored. Note also that, while we have shown how false signals could propagate, our model does not predict any amplification effect. For such an effect to occur, one would have to make additional assumptions, for example, a feedback from the increasing price of property to the assumed profitability of property loans.

5. The Effects of Entry

This section considers a two-period version of the model, where banks commit to a particular minimum level of resources at the outset, believing that both periods will be the same. However, entry occurs during before the second lending period, and we wish to examine the response of the old banks to entry.

The story here is somewhat simpler than in the previous section. Looking first at the identical banks’ case, the entry of some new banks after the first period alters the equilibrium conditions Eq. (7) and Eq. (8), lowering the margins available in both markets. This potentially drives banks into a failure regime where they are optimizing on the deposit-put.

There is a contrast between the situation of a new entry and optimization de novo with the larger group of banks. In the equilibrium with a larger number of banks,
Table 1
Volume of lending and interest rate before and after entry

<table>
<thead>
<tr>
<th>Entrant’s costs</th>
<th>Number of banks</th>
<th>Aggregate lending</th>
<th>Interest rate (low-risk)</th>
<th>Lerner index</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) High profitability</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1.000</td>
<td>.111</td>
<td>.100</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.072</td>
<td>.109</td>
<td>.080</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.118</td>
<td>.107</td>
<td>.067</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.130</td>
<td>.109</td>
<td>.080</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.173</td>
<td>.109</td>
<td>.080</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.202</td>
<td>.105</td>
<td>.067</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.274</td>
<td>.102</td>
<td>.067</td>
<td></td>
</tr>
<tr>
<td>b) Moderate profitability</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1.000</td>
<td>.1048</td>
<td>.045</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.071</td>
<td>.1038</td>
<td>.036</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.114</td>
<td>.1032</td>
<td>.030</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.173</td>
<td>.1024</td>
<td>.030</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.460</td>
<td>.0984</td>
<td>.030</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.765</td>
<td>.0945</td>
<td>.030</td>
<td></td>
</tr>
<tr>
<td>c) Low profitability</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1.000</td>
<td>.1025</td>
<td>.024</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.068</td>
<td>.1020</td>
<td>.019</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.119</td>
<td>.1016</td>
<td>.016</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.212</td>
<td>.1009</td>
<td>.016</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>2.017</td>
<td>.0954</td>
<td>.016</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2.864</td>
<td>.0898</td>
<td>.016</td>
<td></td>
</tr>
</tbody>
</table>

each decides its decision with knowledge of the presence of others, and each therefore curtails its volume of lending. But the unanticipated arrival of a newcomer leaves the existing banks with excess capacity that they will nevertheless employ, given that they have already committed to the resources, which they now treat as sunk costs. The newcomer will not lend as much as it would have had it been in at the start, but (since the others do not contract) its increment will bring total lending above the level it would have reached had all joined at the same time, with a greater risk of driving the old banks into the failure zone.

A further twist occurs if the new entrant has lower unit costs than the incumbents. Depending on the pre-existing profit margins, a new entrant with lower cost could optimally choose a very large scale of operations, driving the old banks into a loss-making position, even on the low-risk loans. (In practice, the cost story would likely be more complex. In order to respond to the entrant, the incumbent might have to increase its own cost base, bidding more aggressively for funds and borrowers, to retain skilled staff, for advertising and so on. This kind of pressure might make the incumbent vulnerable to acquisition, a possibility not modeled here).

Table 1 illustrates these points with a numerical example generated by Eqs. (3–6) above, where we assume values for the cost parameters $w_i$, the number of banks, and
the demand functions $f$ and $g$. Specifically, for Table 1 we set $f = g$, (so that $x_i = z_i$) and, assuming a negative exponential demand function, we compute the equilibrium of the system under optimization of the “no-failure zone” type for 3, 4 and 5 banks, and separately for the equilibrium after a fifth bank enters a four-bank system with sunk costs. The three panels are for different levels of pre-entry profitability (Lerner index), determined by the gap between assumed unit costs and the intercept of the demand function. For the new-entry case, we show three different sub-cases: the first of these has the entrant’s costs at the same level as the incumbents (0.10); the other two sub-cases have lower entrant costs.

Note the large market share taken by the new entrant, and the corresponding sharp decline in the low-risk lending rate, especially when the entrant has low costs and the old profit margin (Lerner index) was low.

These calculations are based on “no-failure zone” optimization. If the old banks move into the failure zone, then the previous discussion showing the likelihood (Condition R) of a shift towards high-risk lending applies.

If, on the other hand, the entry does not move the old banks into the failure zone, what can be said about the impact on the mix of low- and high-risk loans? In the present model, this hinges entirely on the behavior of the two elasticities $\epsilon$ and $\eta$. If these are the same, then the share of high-risk assets will not change with the overall expansion of loan business.

Thus, from the equilibrium conditions Eqs. (5) and (6), where the $w_i$ for the old banks are understood to be the new opportunity cost of the factor resources, and where $H$ is the Herfindahl index, we deduce that

$$\frac{r^L}{\theta r^H} = \frac{1 + \eta H}{1 + \epsilon H}$$

It can be deduced that if $\theta r^H > r^L$ (equivalently $-\eta > -\epsilon$), the difference increases with rising values of $H$. New entry will normally lower the Herfindahl index. Accordingly, in this plausible case, where the expected return on high-risk loans declines more rapidly with aggregate loan volume than does the equilibrium return on low-risk loans, the new equilibrium will involve a lower expected return for high-risk loans, implying a relative shift in lending towards high-risk loans.

In the extreme long run, the fixed costs incurred by the old banks are no longer relevant. Then the row showing the long-run equilibrium with five banks is the relevant one. It is clear that aggregate lending is much lower than in the rows showing the impact effect of the fifth bank entering. It is even possible that the system could return to the no-failure zone. These calculations thus illustrate how the disturbing effect of entry may be larger in the short run.

6. Concluding Remarks

Using a simple model of imperfect competition, we have shown how informational disturbances and entry can upset the equilibrium of a banking system, creating potential
failure problems throughout the system. New entry tends to increase risk-taking and vulnerability to failure of the old banks, especially if they are less cost-efficient than the entrant. Optimistic signals can propagate through the system leading to increased risk of failure even where banks are not exploiting the option value provided them by deposit insurance.

The model provides a coherent theoretical underpinning for much of the recent policy discussion of systemic failure in banking. It can readily be applied or reinterpreted for other disturbances, such as the impact of an insolvent but not yet illiquid bank, or the consequences of taxation of bank intermediation.

Though the model is too simplified to be directly applicable to empirical testing, there are several striking implications that could be borne in mind for empirical work. For example:

- The moderating influence on risk-taking of franchise value measured, as has been suggested above, by the gross return on low-risk lending could potentially be identified in cross-section analyses of banking systems in different countries, or in a time series sufficiently long to show some variance in franchise value thus measured.
- The contagious flow of information through the different participants in the banking market could be empirically analyzed given bank-by-bank data on the share of lending to higher risk areas such as commercial property.
- The dynamic pattern, whereby the predicted increase in risk behavior following new entry is only temporary, appears to be novel, and could be tested.

We find that a high franchise value for low-risk banking may keep banks away from the failure zone, and even if it is not high enough to do that, it may prevent the bank’s behavioral response to failure zone behavior from increasing risk. This is consistent with standard prescriptions that call for high capitalization, and also with the idea that limiting entry reduces the risk of bank failure. However, the discussion has not been explicitly couched within a policy framework; therefore it would not be wise to jump to over-hasty policy conclusions. Even if bank entry can destabilize existing banks, that is not a sufficient reason to block entry. Other solutions can be considered: strong supervision may counter the tendencies noted and entry through acquisition may be less disruptive.

Acknowledgments

This research was commenced during a stimulating stay at the Research Department of the International Monetary Fund, for whose hospitality I am grateful.

Notes

2. Reference should also be made to the many models of individual bank behavior, e.g. Guttentag and Herring (1986) and Keeley (1990), to mention two contrasting contributions. A different strand is the macroeconomic financial crisis literature, e.g. Fisher (1933), Mankiw (1986) and Minsky (1995). Models of noise transmission in securities markets have been developed, e.g. by de Long et al. (1990a,b).

3. Clearly, there would be some variation in the cost of capital with the risk assumed; this is neglected here for simplicity. Unless the required capital is very high, the variation would be of second-order importance for determining equilibrium behavior.

4. Insofar as the banks can tell whether a loan is low- or high-risk, we are abstracting from the adverse selection issue stressed in the literature, following Stiglitz and Weiss (1981).

5. In reality, no doubt, such costs are higher for high-risk loans, but we retain equal cost for simplicity. (We could also add a fixed cost component without altering the marginal conditions.)

6. Reflecting the existence of only one distinguishable risk-category.

7. For other applications of the Cournot model to banking see Honohan and Kinsella (1982), Steinherr and Huveneers (1994) and Freixas and Rochet (1997, pp. 59–61).

8. The assumption of non-diversifiability is essential, but the dichotomous pattern of returns is made for simplicity, and could be modified at the expense of additional complexity. For example, we could alternatively assume that the stochastic amount repaid on the risky portfolio was uniformly distributed on 

9. Bearing in mind the dependence of the rate of return on total lending through the functions $f$ and $g$.

10. The condition can be expressed as:

$$\frac{r^H - (r^* + w_i)}{r^L - b_i} > \frac{\theta r^H}{r^* + w_i} x_i Z \eta$$

For low capital requirements, the left hand side is the ratio of the profitability of repaid high-risk loans to what we have called the franchise value of low-risk banking; the left hand side is the average profit rate on high risk loans, multiplied by the market share multiplied by the elasticity.

11. This kind of updating is a special case of those discussed by Sargent (1993).

12. Specifically: $r^* = 1 - \alpha \exp(-\beta X)$. The results are not qualitatively very sensitive to different values of the parameter $\beta$, as the equilibrium responds to changing $\beta$ by moving along the demand curve to where the elasticity is the same.
13. The parameters are $\beta = 0.3; r^* + w - 1 = 0.1$ (cost for old banks); $\alpha = 0.15$ (high profitability), 0.12 (medium profitability), and 0.11 (low profitability). The Lerner index in the table is expressed as a proportion of interest rate, not of gross interest factor as in the theoretical development.

References


