



FEDERAL RESERVE BANK OF PHILADELPHIA

Ten Independence Mall
Philadelphia, Pennsylvania 19106-1574
(215) 574-6428, www.phil.frb.org

Working Papers

Research Department

WORKING PAPER NO. 97-25

THE WINNER'S CURSE IN BANKING

Sherrill Shaffer

November 1997

WORKING PAPER NO. 97-25

THE WINNER'S CURSE IN BANKING

Sherrill Shaffer

November 1997

Much of this paper was completed while the author was affiliated with the Federal Reserve Bank of Philadelphia. This paper represents the views of the author and does not necessarily represent the views of the Federal Reserve Bank of Philadelphia or of the Federal Reserve System. The author is grateful for research assistance from Louise Berna.

ABSTRACT

Theoretical studies have noted that loan applicants rejected by one bank can apply at another bank, systematically worsening the pool of applicants faced by all banks. This paper presents the first empirical evidence of this effect and explores some additional ramifications, including the role of common filters, such as commercially available credit scoring models, in mitigating this adverse selection, implications for de novo banks, implications for banks' incentives to comply with fair lending laws, and macroeconomic effects.

JEL codes: G2, L1

Sherrill Shaffer
Department of Economics and Finance
University of Wyoming
P.O. Box 3985
Laramie, WY 82071-3985
(307) 766-2173
Fax: (307) 766-5090
e-mail: shaffer@uwyo.edu

THE WINNER'S CURSE IN BANKING

1. Introduction

Theoretical studies by Broecker (1990) and Nakamura (1993) have identified a winner's curse in bank lending, resulting from the ability of rejected loan applicants to apply at additional banks. To the extent that credit screening is imperfectly correlated across banks, poorer credit risks thereby gain a greater chance of being misidentified as good risks, as an increasing function of the number of banks in the market.

No study thus far has explored empirical evidence for this effect. Moreover, at least three theoretical aspects of this effect have also been previously overlooked. First, to the extent that banks rely on common filters (shared databases and uniform screening criteria) in assessing applications, the adverse selection can be theoretically mitigated. An immediate corollary is that the use of credit bureau databases and commercially standardized credit scoring models would have the potential to provide this sort of improvement, even if the credit scoring models are somewhat less accurate than bank-specific traditional credit analysis. A further corollary is that banks may be more likely to adopt such standardized scoring models in unconcentrated geographic or product markets and during recessions.

Second, de novo banks (recent entrants) should be particularly susceptible to adverse selection, as any new bank in a community will face a pool of potential applicants that includes a backlog of those previously rejected over some period of time, in addition to the normal flow of recent rejects encountered by mature banks in a market. This effect is generally recognized by practitioners, but has never been formally quantified. Because roughly 3,000 new banks have

formed within the past 15 years in the U.S. despite substantial net structural consolidation, the aggregate impact of this problem is potentially large.

Third, although previous studies have focused on the case in which some correlates of loan performance are unobservable to the bank, rendering the credit screening process noisy, the same model would also characterize outcomes when banks are prohibited by law from conditioning their decisions on one or more observable correlates of loan performance. Under certain assumptions that have been empirically supported, this could mean that banks have a profit incentive to resist particular forms of consumer legislation in concentrated markets more than in unconcentrated markets.

Besides elaborating on this richer theoretical context of adverse selection, the present study develops two forms of empirical evidence of the winner's curse in banking and explores some indications of its broader macroeconomic effects.

2. Theoretical Considerations

Consider a market with N potential borrowers per period and n banks. Borrowers are of two types as in Broecker (1990), Ferguson and Peters (1995), and Shaffer (1996). "Good" borrowers repay a fraction θ_H of their loans on average while "bad" borrowers repay a fraction θ_L of their loans on average, where $1 > \theta_H$ ("high") $> \theta_L$ ("low") > 0 . A fraction α of these N potential borrowers are good while the remainder $(1 - \alpha)$ are bad. Banks lend at an interest rate r that is fixed, as in Ferguson and Peters (1995) and Shaffer (1996); and exogenous, as in Nakamura (1993). A bank earns an expected profit of $r\theta_H - 1 + \theta_H$ on each \$1 loan made to a good borrower and $r\theta_L - 1 + \theta_L$ on each \$1 loan made to a bad borrower.¹ To ensure a sustainable equilibrium, we assume $r\theta_H - 1 + \theta_H > 0$. To

rule out the trivial equilibrium in which no bank would bother to screen applicants, we assume $r\theta_L - 1 + \theta_L < 0$.

When a potential borrower applies at a bank, that bank observes a signal of the borrower's type and extends the loan if and only if the signal indicates that the borrower's type is "good" (H). Signals observed by multiple banks for a given borrower are assumed to be i.i.d., where each bank infers the correct type with probability p_{hH} for good borrowers and p_{lL} for bad borrowers, and incorrectly infers the type with probability $1 - p_{hH}$ for good borrowers and $1 - p_{lL}$ for bad borrowers. Because a perfectly random or uninformative signal would have $p_{hH} = p_{lL} = 1/2$, we assume $1/2 < p_{hH} < 1$ and $1/2 < p_{lL} < 1$ so that signals are noisy but somewhat informative. Each signal can be interpreted as the outcome of the bank's credit screening process, whether a traditional loan analysis or an automated credit scoring model.

If there is only one bank in the market, all potential borrowers apply at that bank. The bank deems $p_{hH}\alpha N + (1 - p_{lL})(1 - \alpha)N$ applicants to be creditworthy and makes this many loans for an expected profit of $p_{hH}\alpha N(r\theta_H - 1 + \theta_H) + (1 - p_{lL})(1 - \alpha)N(r\theta_L - 1 + \theta_L)$. The expected loan loss rate for this bank is $[p_{hH}\alpha(1 - \theta_H) + (1 - p_{lL})(1 - \alpha)(1 - \theta_L)] / [p_{hH}\alpha + (1 - p_{lL})(1 - \alpha)]$. Rejected applicants cannot subsequently reapply.

If two banks occupy the market, we assume that half of the potential borrowers initially apply to one bank, and the other half to the other bank. This symmetric case corresponds to Nakamura's "anonymous lender" scenario. Any applicant rejected by one bank can subsequently apply at the other bank. Thus, each bank makes half as many loans to initial applicants as would the monopoly bank and earns half the expected profit from this subset of borrowers, but additionally faces $(1 - p_{hH})\alpha N/2 + p_{lL}(1 - \alpha)N/2$ applicants who were previously rejected by the other bank. A crucial point

is that the banks cannot distinguish between initial applicants (those who have not been previously rejected by another bank) and subsequent applicants (those who have been previously rejected by another bank). Therefore, in addition to its loans made to a subset of initial applicants, each bank makes $p_{hH}(1 - p_{hH})\alpha N/2 + (1 - p_{lL})p_{lL}(1 - \alpha)N/2$ loans to second-round applicants previously rejected by the other bank. The total number of loans made by both banks equals $p_{hH}\alpha N(2 - p_{hH}) + (1 - p_{lL})(1 + p_{lL})(1 - \alpha)N$, which exceeds that made by the monopoly bank since $2 - p_{hH} > 1$ and $1 + p_{lL} > 1$. Expected profits for each bank are $p_{hH}\alpha N(2 - p_{hH})(r\theta_H - 1 + \theta_H)/2 + (1 - p_{lL})(1 + p_{lL})(1 - \alpha)N(r\theta_L - 1 + \theta_L)/2$, and the expected loan loss rate for each bank (and for the banking industry in this market) is $[p_{hH}\alpha(2 - p_{hH})(1 - \theta_H) + (1 - p_{lL})(1 + p_{lL})(1 - \alpha)(1 - \theta_L)] / [p_{hH}\alpha(2 - p_{hH}) + (1 - p_{lL})(1 + p_{lL})(1 - \alpha)]$.

In general, where a market contains n identical banks and potential borrowers apply at randomly selected banks until they either receive a loan or have been rejected by all n banks, each bank receives a total number of applicants per period given by:

$$(1) \quad \text{Applicants per bank} = (\alpha N/n) \sum_{m=1}^n (1 - p_{hH})^{m-1} + (1 - \alpha)(N/n) \sum_{m=1}^n p_{lL}^{m-1}$$

and makes loans to those its signals indicate to be creditworthy:

$$(2) \quad \text{Loans per bank} = (p_{hH}\alpha N/n) \sum_{m=1}^n (1 - p_{hH})^{m-1} + (1 - p_{lL})(1 - \alpha)(N/n) \sum_{m=1}^n p_{lL}^{m-1}.$$

The aggregate number of loans made to the fixed pool of N potential borrowers is:

$$(3) \quad \text{Total loans} = p_{hH}\alpha N \sum_{m=1}^n (1 - p_{hH})^{m-1} + (1 - p_{lL})(1 - \alpha)N \sum_{m=1}^n p_{lL}^{m-1}$$

which is an increasing function of n since $(1 - p_{hH})^{m-1} > 0$ and $p_{lL}^{m-1} > 0$ for all $m \geq 1$. That is, the more banks inhabit the market, the fewer potential borrowers are ultimately unable to obtain a loan. Because $p_{hH} > 1/2 > 1 - p_{lL}$, a higher proportion of good applicants are accepted than the proportion of accepted bad applicants in each round, with the result that each successive pool of rejected applicants contains a higher proportion of bad risks than the previous round. This worsening of the pool, previously noted by Broecker (1990) and Nakamura (1993) and analyzed in more detail below, generates an expected loan loss rate that is an increasing function of the number of banks:

(4) Expected loss rate =

$$\frac{[p_{hH}\alpha(1 - \theta_H)\sum_{m=1}^n (1 - p_{hH})^{m-1} + (1 - p_{lL})(1 - \alpha)(1 - \theta_L)\sum_{m=1}^n p_{lL}^{m-1}]}{[p_{hH}\alpha\sum_{m=1}^n (1 - p_{hH})^{m-1} + (1 - p_{lL})(1 - \alpha)\sum_{m=1}^n p_{lL}^{m-1}]}$$

$$= \{\alpha(1 - \theta_H)[1 - (1 - p_{hH})^n] + (1 - \alpha)(1 - \theta_L)(1 - p_{lL}^n)\} / \{\alpha[1 - (1 - p_{hH})^n] + (1 - \alpha)(1 - p_{lL}^n)\}.$$

The change in the expected loss rate, comparing a market with n banks to one with $n - 1$ banks, reduces after some algebra to:

(5) Change in expected loss rate, n banks versus $(n - 1)$ banks =

$$p_{hH}(1 - p_{lL})\alpha(1 - \alpha)(\theta_H - \theta_L)[p_{lL}^{n-1} \sum_{m=1}^{n-1} (1 - p_{hH})^{m-1} - (1 - p_{hH})^{n-1} \sum_{m=1}^{n-1} p_{lL}^{m-1}]$$

which has the sign of the expression in square brackets. Comparing this latter expression term by term, and recalling that $p_{iL} > \frac{1}{2} > (1 - p_{hH})$, we note that $p_{iL}^a(1 - p_{hH})^b - (1 - p_{hH})^a p_{iL}^b = p_{iL}^b(1 - p_{hH})^b [p_{iL}^{a-b} - (1 - p_{hH})^{a-b}] > 0$ for all $a - b > 0$. Substituting $n - 1$ for a , and each successive value in the summations of $m - 1$ for b , we obtain the key result that *an increase in the number of banks monotonically increases the expected loan loss rate.*

Expected profits for each bank i are:

$$\begin{aligned}
 (6) \quad E\pi_i &= p_{hH}\alpha(N/n)(r\theta_H - 1 + \theta_H)\sum_{m=1}^n (1 - p_{hH})^{m-1} \\
 &\quad + (1 - p_{iL})(1 - \alpha)(N/n)(r\theta_L - 1 + \theta_L)\sum_{m=1}^n p_{iL}^{m-1} \\
 &= \alpha(N/n)(r\theta_H - 1 + \theta_H)[1 - (1 - p_{hH})^n] + (1 - \alpha)(N/n)(r\theta_L - 1 + \theta_L)(1 - p_{iL}^n).
 \end{aligned}$$

Since $1 - (1 - p_{hH})^n > 1 - (1 - p_{hH})^{n-1}$ for all $n > 1$, the contribution of good borrowers to expected profits is an increasing function of n . At the same time, since $1 - p_{iL}^n > 1 - p_{iL}^{n-1}$ for all $n > 1$ and $(r\theta_L - 1 + \theta_L) < 0$ by assumption, the drain on profits by bad borrowers is also greater for larger n . We can assess how these contrary effects net out across the entire banking market by calculating the change in aggregate expected profits, comparing a market with n banks to one with $n - 1$ banks:

$$(7) \quad \Delta E(\sum \pi_i) = p_{hH}(1 - p_{hH})^{n-1}\alpha N(r\theta_H - 1 + \theta_H) + p_{iL}^{n-1}(1 - p_{iL})(1 - \alpha)N(r\theta_L - 1 + \theta_L).$$

In attempting to sign this expression, note that $1 > p_{hH} > \frac{1}{2}$ and $1 > p_{iL} > \frac{1}{2}$ implies that $p_{hH}(1 - p_{hH})^{n-1} < p_{iL}^{n-1}(1 - p_{iL})$ for sufficiently large n . Also, $(r\theta_H - 1 + \theta_H) > 0$ while $(r\theta_L - 1 + \theta_L) < 0$ by assumption.

Even so, the sign of equation (7) depends on the parameter values. Thus, while the loan loss rate increases with the number of banks, it is possible for expected aggregate bank profits to increase with n if there are enough type H borrowers and if each type H loan is sufficiently profitable.²

Applicant Attrition

A further detail of adverse selection among bank borrowers concerns "applicant attrition." A rational potential borrower will apply for a loan as long as the expected benefit of applying equals or exceeds the cost of applying. The expected benefit equals the probability of receiving the loan times the benefit of the loan. The cost of applying includes both direct costs such as application fees and other transaction costs, and indirect costs or opportunity costs resulting from the time needed to obtain and file each application. The benefits of some loans may be time-dependent and decline with any delays in granting the loan, as with seasonal agricultural loans, construction loans at different phases of the business cycle, or a home mortgage where the buyer may lose the house if financing is not approved by some deadline. Extensive delays in granting certain business loans may impair the borrower's financial condition and ability to repay, further undermining her ability to qualify for a loan. A succession of rejections may also lead an applicant to revise downward her perceived likelihood of being granted a loan on subsequent applications. For all these reasons, the equilibrium number of applications per borrower for a given loan should be finite and an applicant who has been rejected several times will eventually stop applying at additional banks.

The maximum number of applications in equilibrium will vary across borrowers, but the overall impact of such considerations implies that the linkage between bank structure and average applicant quality will be nonlinear. In particular, the magnitude of the *marginal* effect of the number

of banks on average credit quality should be a decreasing function of the number of banks, approaching zero in the limit for markets containing more banks than the maximum equilibrium number of applications for a given potential borrower. This nonlinearity should be exacerbated by the finiteness of available financial information about a given applicant, which implies that banks' "draws" may not be completely independent of each other and, beyond some number of banks, may constitute convex combinations of prior draws. Thus, once an applicant has been rejected a certain number of times, the likelihood of subsequent acceptance approaches zero. Such nonlinearity is tested in the empirical section below by substituting for the number of banks a concave function (the logarithm) thereof.

Common Filters

The adverse selection characterized above relies on each bank's observing separate, imperfectly correlated signals about a given borrower's type. Clearly, if all banks observe the same signals, an applicant rejected by one bank would be rejected by all and market structure would have no impact on the loan loss rate or on the number of loans granted in aggregate. Certain types of bank loans appear to resemble this situation conceptually. Credit analysis of consumer loans (including credit cards) typically involves a background credit check of widely accessible databases maintained by one of three large consumer credit bureaus (Equifax, Trans Union, or Experian/TRW). Thus, every potential lender has access to the same information about a given applicant, a factor that would tend to increase the correlation of perceived creditworthiness across banks. In addition, standardized credit scoring models are now commercially available for credit cards, mortgage lending, and small business lending (see Altman and Haldeman, 1995; Asch, 1995; Avery et al., 1996), with some (such

as the Fair, Isaac--or FICO--credit score) actually maintained on the databases of the credit bureaus. Combining common information sets with common selection criteria can further increase the correlation of credit analysis and, potentially, of lending decisions across banks.

To the extent that adverse borrower selection is a problem in unconcentrated banking markets, credit bureaus and standardized credit scoring models can potentially mitigate the severity of this problem. Conversely, to the extent that adverse borrower selection is among the various factors driving the recent and ongoing wave of consolidation among U.S. banks, the further development of improved and widely accepted credit scoring models might somewhat reduce the incentive toward additional consolidation.³

Statistical scoring models have been developed in a manner ensuring that, on a historical basis, they are more accurate than traditional credit analysis using the same data. Of course, past performance is no guarantee of future accuracy. The most widely available of these models use no institution-specific data, and all are limited by the degree of accuracy and timeliness of the databases on which they draw. Custom scoring models, while capable of using bank-specific data, are often limited by small samples. Consequently, controversy has persisted regarding the overall accuracy of automated credit scoring compared with traditional methods.

An important point in this regard is that, according to the model above, banks in unconcentrated markets would suffer a higher loan loss rate using idiosyncratic filters than using an equally accurate common filter. Therefore, a common filter may even be somewhat less accurate than traditional, independently applied financial analysis, and still achieve a lower loan loss rate. The more banks operate in a market, the more the benefits of a common filter--by circumventing the

problem of repeated applications by previously rejected applicants--can offset some intrinsic loss of accuracy in the screening process.

Thus, the rational stance regarding commercially available credit scoring models versus traditional methods may vary systematically across banking markets. A local monopoly bank would prefer to use the FICO score, for example, only if it is consistently more accurate in identifying borrower types than the bank's in-house analysis. Each bank in a large urban market with dozens of rival lenders may do better to use the FICO score, by contrast, even if it is not the most accurate one attainable, simply in order to achieve a uniform outcome across banks.⁴

To quantify this effect, let β_{hH} and β_{lL} denote the probabilities of accurately identifying good and bad borrowers, respectively, by means of the common filter. Then, if all banks in the market rely on the common filter, the mean loan loss rate would be $[\beta_{hH}\alpha(1 - \theta_H) + (1 - \beta_{lL})(1 - \alpha)(1 - \theta_L)]/[\beta_{hH}\alpha + (1 - \beta_{lL})(1 - \alpha)]$. The minimum (i.e., least accurate) values of β_{hH} and β_{lL} that would equate this loan loss rate with that given by equation (3) above are given by:

$$(8) \quad \beta_{hH}/(1 - \beta_{lL}) = [1 - (1 - p_{hH})^n]/(1 - p_{lL}^n)$$

which is less than $p_{hH}/(1 - p_{lL})$ since $1 - p_{hH} < 1/2 < p_{lL}$. Note that this expression is independent of α , θ_H , and θ_L .

The magnitude of this effect can be quite striking. If $p_{hH} = p_{lL} = 0.9$ and $n = 20$, then condition (8) is satisfied by $\beta_{hH} = \beta_{lL} \approx 0.532$. That is, for only 20 banks in the market, a common filter that is only slightly better than random can achieve the same loan loss rate as an i.i.d. filter that is 90 percent accurate.

Of course, banks' choice between an i.i.d. filter and a common filter will depend on expected profits, which is only indirectly related to the loan loss rate. The condition for a common filter to generate the same expected profits as i.i.d. signals is:

$$(9) \quad \beta_{\text{HH}}\alpha(r\theta_{\text{H}} - 1 + \theta_{\text{H}}) + (1 - \beta_{\text{LL}})(1 - \alpha)(r\theta_{\text{L}} - 1 + \theta_{\text{L}}) = \\ [1 - (1 - p_{\text{HH}})^n]\alpha(r\theta_{\text{H}} - 1 + \theta_{\text{H}}) + (1 - p_{\text{LL}})^n(1 - \alpha)(r\theta_{\text{L}} - 1 + \theta_{\text{L}}).$$

Because this expression contains all the parameters in the model, no clear qualitative conclusion emerges, so it is instructive to consider two numerical examples. As above, let $p_{\text{HH}} = p_{\text{LL}} = 0.9$ and $n = 20$. Also let $\theta_{\text{H}} = 1$, $\theta_{\text{L}} = 0.5$, $\alpha = 0.8$, and $r = 0.1$. Then the right-hand side of equation (9) equals 0.000942 so that the condition can be satisfied by $\beta_{\text{HH}} = \beta_{\text{LL}} \approx 0.535$. In this case, banks will prefer the common filter even if its intrinsic accuracy is substantially worse than that of the i.i.d. filter and indeed little better than random. The common filter has a similar effect on expected profits as on expected loan loss rates for these parameter values.

By contrast, consider the same parameter values except for $\theta_{\text{L}} = 0.8$ and $\alpha = 0.9$. Then the right-hand side of equation (9) is about 0.0795 and the condition is satisfied by $\beta_{\text{HH}} = \beta_{\text{LL}} \approx 0.897$. Thus, with a higher proportion of good borrowers and better performance of bad borrowers, it is possible for banks to prefer the i.i.d. signals even when the common filter is nearly as accurate, despite the common filter's ability to reduce expected loan loss rates as shown above. Together, these two numerical examples suggest that banks' preference for a common filter will be stronger in a recession, when defaults are more common, than during the expansion phase of a business cycle.

The rational outcome may also vary systematically across types of bank loans, depending on the relevant geographic market for each. For instance, the geographic market for credit card loans is essentially nationwide, since there is no need for the lender to be familiar with local economic conditions or local collateral values. For such products, unless a bank is able somehow to differentiate its product from those of its rivals, the effective number of competitors on the lending side is larger than for more geographically localized products, leading to both a stronger adverse selection effect and a stronger incentive to adopt common filters. Our model implies that this may be one reason why credit card lending has been among the types of bank services for which the adoption of credit scoring models has been particularly widespread.

Of course, banks can customize their *implementation* of the FICO score or equivalent, with each bank selecting a different acceptance threshold or using different ranges of the score to apply different collateral, pricing, or guarantee requirements. In such cases, borrowers may learn that some banks are more willing than others to lend to riskier applicants, and this learning process might initially entail several rounds of iterated applications. Nevertheless, the essential point is that the *signals* generated by a common filter are perfectly correlated across banks, even if banks' responses to the signals vary systematically.⁵ Given a stable set of policies across banks, one expects that borrowers would eventually begin to self-select in a sequential or hierarchical fashion, with the most credit worthy borrowers applying at banks setting the highest acceptance threshold, and the least credit worthy applying to those banks setting the lowest acceptance threshold. Such a self-selection equilibrium, which is beyond the scope of this paper to analyze formally, would entail risk-based pricing of loans such that the most selective banks charge the lowest rate (thereby attracting the most

creditworthy borrowers) while the least selective banks charge the highest interest rates to cover their higher loan loss rates. In this way the self-selection process could be sustained as an equilibrium.

De Novo Banks

In a market with a fixed number of banks, a backlog will accrue over time of applicants who have been rejected by every bank, subject to the considerations of borrower attrition discussed above. If a new bank then enters, this backlog of rejected applicants will apply at the new bank along with the usual mix of first, second, and subsequent applicants common to all banks in the market. The fraction of the de novo bank's total applicant pool reflecting this backlog will depend on the maximum number of periods following the original rejection that such an applicant is willing to continue applying. While that number may be variable, we can at least characterize the mix of types within this backlog and compare that mix with the proportions of types facing mature banks in the same market. It was previously noted in the context of loan loss rates that the pool of applicants worsens with each successive round of rejection; here we quantify explicitly the mix of applicants and the change in that mix as a function of the number of rejections.

The mix of applicant types, or ratio of good to bad applicants, within the pool of those who have been rejected n times is just $\alpha(1 - p_{hH})^n / [(1 - \alpha)p_{\ell L}^n]$. The corresponding ratio for those who have been rejected $n - 1$ times is $\alpha(1 - p_{hH})^{n-1} / [(1 - \alpha)p_{\ell L}^{n-1}]$. For any n , the ratio of these two ratios is $(1 - p_{hH}) / p_{\ell L}$, which is less than 1 since $p_{hH} > 1/2$ and $p_{\ell L} > 1/2$. Thus, at each successive stage of rejection, the remaining pool of potential borrowers going forward to the next round of applications is monotonically worse (i.e., has a lower proportion of good applicants versus bad applicants). For any value of n , therefore, the backlog facing a new bank entering the market has a worse mix and will

generate higher loan loss rates than the group of potential borrowers still applying among the incumbent banks. As long as the entrant bank uses a similar loan screening technology with accuracy similar to that of the technology employed by the incumbent banks, the de novo bank will suffer a higher loan loss rate than the mature incumbents. This prediction will be tested in the empirical section below.

Adverse Selection and Fair Lending

An important and long-recognized social issue concerns the possibility that some banks may condition their credit analysis on observable characteristics that have no direct bearing on the individual applicant's ability or willingness to repay, yet have historically correlated with default rates in excess of what can be predicted based solely on other observable, strictly financial factors. In response to this problem, U.S. federal law prohibits lenders from conditioning their credit analysis and lending decisions on certain nonfinancial factors such as race, gender, or age. The model above can suggest new implications for the linkage between market structure and the willingness of lenders to comply with such fair lending laws.

In particular, suppose now that the hypothesized signals represent information contained in one of the prohibited criteria. By law, banks must ignore such signals and lend on an equal basis to all.⁶ The general result here is that, because the screening process (were it permitted to be used) is less effective in less concentrated markets, banks in those markets will have less financial incentive to oppose or circumvent the fair lending laws than would banks in more concentrated markets.

3. Empirical Findings

This section tests some of the theoretical implications discussed above, along with others developed below. First, we examine loan chargeoff ratios as a function of the age of the bank in a fixed-effects linear regression model containing quarterly calendar time dummies as well as annual age-of-bank dummies. The sample spans all U.S. commercial banks during 1986-1995, a quarterly sample containing nearly half a million observations. Bankers, regulators, and academic researchers have consistently considered that banks exhibit atypical performance during their first five years, so we are particularly interested in quantifying this pattern, as well as exploring the extent to which atypical chargeoff ratios may extend somewhat beyond the five-year horizon.⁷ We included age dummies for each of a bank's first 10 years, which is longer than the period tracked by most previous studies of de novo bank performance and longer than observed periods of anomalous profitability; see DeYoung and Hasan (1997).

Consistent with bankers' common perception, net chargeoff rates, as shown in Table 1, are significantly below average during the first year of a bank's existence and about average during the second year. In years 3 through 6, the average chargeoff rate is roughly double that of mature banks. In years 7 through 9, the chargeoff rate is still about 40 - 50 percent higher than that of mature banks, with the difference statistically significant. Only in a bank's tenth year does the chargeoff experience approach that of more mature banks. These findings not only confirm bankers' perceptions and support the theory, but they also show that transitional effects of bank entry persist nearly twice as long as regulators and academic researchers have believed.

We must consider whether our findings reflect actual adverse selection of loan applicants, or merely a learning process involving inexperienced lenders. De novo banks always begin operation with experienced staff and management in selected key positions, as this is a regulatory

prerequisite to obtaining a bank charter. Thus, it is unlikely that the findings reflect merely a learning process. However, even if the predictions of the adverse selection model are "right for the wrong reason," the empirical results remain instructive in their own right.

As a test of robustness, and to explore systematic differences in recovery rates, the same regression was run for gross chargeoff rates. Table 2 shows that banks experience significantly lower gross chargeoff rates than mature banks during their first two years and significantly higher gross chargeoff rates in years 3 through 9, a pattern of age dependency similar to that reported in Table 1. Comparing the right-hand columns of Tables 1 and 2, we see that the magnitude of the excess is smaller for gross chargeoff rates than for net chargeoff rates during these latter years, implying that recovery rates as a fraction of gross chargeoffs are lower for de novo banks than for mature banks on average.

These results are consistent with, and complementary to, results on the profit efficiency of de novo banks reported by DeYoung and Hasan (1997). That study found that new banks are significantly less profitable than older banks of the same asset size until the ninth year of a bank's operation. However, the pattern of loan delinquencies observed in that study did not appear to be a contributing factor to the lower profitability, and loan chargeoffs were not explored. DeYoung and Hasan included age-of-bank dummies for the first 14 years of a bank's charter, unlike previous studies of post-deregulation de novos, which quantified the impact of only the first eight years.

A second empirical approach focuses on mature banks, each operating in a single geographic market (MSA). Using a cross-sectional sample of nearly 3,000 banks in over 300 MSAs across the U.S. as of year-end 1990, we find that loan chargeoff rates are a significantly increasing function of the total number of banks in the MSA, consistent with the theory (Table 3).⁸ This relationship

persists across total loans even when the population of the MSA (not shown) and other characteristics are controlled for. The estimated magnitude of the effect implies that each additional rival bank drives up the gross chargeoff rate of each incumbent by 0.10 basis points, or 0.10 percent of the sample mean chargeoff rate. Since the number of incumbent banks ranges from 4 to 311 across the sample MSAs, the aggregate impact on loan chargeoffs is economically significant.

Table 3 also shows that the log of the number of banks is even more significant and yields a higher adjusted R-square than the simple number of banks, implying a nonlinear relationship between bank structure and performance consistent with the theoretical considerations spelled out in the previous section. The net chargeoff ratio increases by 2.7 basis points as the number of banks in an MSA increases from five to six, or by 1.6 basis points as the number of banks increases from nine to 10, for example. Thus, for MSAs containing a small number of banks, structure has a greater magnitude of effect on chargeoffs in the semilog model than in the linear model.

Again, it is possible that the empirical linkage between bank structure and ex post credit quality may be driven by other factors besides adverse borrower selection. While the available data cannot distinguish among alternative explanations, the empirical finding has apparently not been previously recognized and merits closer scrutiny.

However, the effect is not uniform across all types of loans. When chargeoffs are disaggregated by loan type, as shown in the last six columns of Table 3, bank structure exhibits a significant effect on commercial loans but an insignificant effect on consumer loans and "other" loans (including real estate and agricultural loans). Possible reasons for these differences might include the role of collateral in real estate and agricultural loans, common data (for example, from credit bureaus) used for consumer loans, standardized lending criteria across certain categories of

real estate loans, and other factors. In addition, the geographic market for some loan types, such as credit card loans, may not be well approximated by a single MSA, rendering our structural measures inappropriate in those cases. Data for savings and loan associations were unavailable, and although the number of S&Ls should be positively correlated with the number of banks in an MSA, it is possible that the exclusion of S&Ls may have weakened the results for those types of loans (such as consumer or residential real estate loans) in which S&Ls compete most strongly with commercial banks. The net chargeoff ratio for commercial loans exhibited a nonlinear response to bank structure, as suggested by theory, since the log of the number of banks was more significant than the simple number of banks.⁹

Economic Growth

A further important question concerns the welfare effects of this apparent adverse selection. Since banks are only imperfectly able to distinguish *ex ante* between good credits and bad credits, perfect risk-based pricing is not possible (and would carry its own adverse selection effects as analyzed by Broecker, 1990, besides being subject to the intrinsic limitations analyzed by Stiglitz and Weiss, 1981). Thus, within a market, good borrowers must somewhat subsidize bad borrowers. The ambiguous effect of the adverse selection on banks' expected profits, shown theoretically above, indicates that the overall welfare effects are not obvious but must be analyzed explicitly.

The broader effect of adverse borrower selection on the economy hinges on whether the primary distinction between creditworthy and uncreditworthy borrowers involves the mean or the dispersion of the distribution of returns on their respective projects. The nature of the loan contract limits the extent to which banks can share in the upper tail of volatile returns; thus, if bad credit risks

face projects that are more volatile than those of good credit risks but have comparable mean returns, then funding those projects could be beneficial to society on balance, even though it may represent a redistribution of expected profits from the banking sector to the real sector. The recently recognized importance of the financial intermediation sector in fostering economic growth (see, for example, Jayaratne and Strahan, 1996; Krol and Svorny, 1996) underscores the relevance of this question.

Regressions addressing this issue are summarized in Table 5 using data characterized in Table 4. The dependent variable is the percentage change in money income per household between 1979 and 1989 by MSA for a nationwide sample, analogous to the statewide economic growth variables explored in Jayaratne and Strahan (1996) and Krol and Svorny (1996). Explanatory variables include the MSA population as of 1980, the percentage of adults in each MSA with four or more years of college education, and the number of banks with one or more offices in each MSA during the years 1979 or 1984. In the linear equations (odd-numbered models), the point estimates correspond to an incremental 10-year income growth of 10 to 13 basis points per additional bank in the MSA. When population is included in the regression, its collinearity with the number of banks ($\rho = 0.89$ for both 1979 and 1984) inflates the standard error of the coefficient on the latter variable. Of the two variables, the number of banks exhibits the more significant association with growth when both are included.

The even-numbered models in Table 5 incorporate the natural logarithm of the number of banks operating in each MSA. This form is intended to capture the anticipated effect of "applicant attrition" as discussed above, assuming that the linkage between structure and growth reflects to some extent the adverse selection of applicants. Comparing these results with those of the odd-

numbered models, we see that the log variables all exhibit larger t-ratios and yield higher adjusted R-squares than the corresponding linear models. This outcome may reflect the mitigation of multicollinearity, since the correlation between the log of the number of banks and population is only 0.67 for both 1979 and 1984. However, these results also support a nonlinear linkage between bank structure and community income growth. The estimated coefficients indicate that an additional bank is associated with an increase of 62 to 72 basis points in the MSA's 10-year growth rate for an MSA with five banks initially, 36 to 41 basis points for an MSA with nine banks, and 14 to 16 basis points for an MSA with 24 banks, for example.

Because of multicollinearity between the number of banks and population, we considered it important to explore the robustness of the estimates by estimating several variants of the model. Some of these variants used subsets of MSAs with populations between 100,000 and 1 million, or between 100,000 and 500,000, as a way of reducing the severity of the multicollinearity. In all cases, the number of banks operating in each MSA in 1979 was positively and significantly associated with the subsequent economic growth rate of that MSA; the magnitude and significance of the corresponding coefficient was greater for the subsamples (not shown in the table) than for the full sample. Similarly, the coefficient on the number of banks in 1984 was significantly positive in nearly all cases.

As further tests of robustness, the models were estimated with additional control variables (not shown in the table) including banks per capita, public expenditures per capita, and the percentage of Democratic votes in the 1992 presidential election (the last two variables suggested by Jayaratne and Strahan, 1996). These variables were not statistically significant and did not materially alter the magnitude or significance of the estimated coefficients on the other variables.

In general, it appears to be a robust finding that a larger number of banks within an MSA is associated with a significantly higher concurrent or subsequent growth rate in per capita income. In the linear models, the stronger association seems to be with subsequent growth rather than contemporaneous growth: the bank numbers for 1979 generated larger adjusted R-squares, generally larger point estimates, and generally larger t-statistics than the bank numbers for 1984 in corresponding models (comparing models 1 and 5, and 3 and 7 in Table 5). In the semilog models, the regressions incorporating 1979 bank structure data exhibited slightly larger adjusted R-squares, although 1984 bank structure figures yielded larger point estimates and t-ratios (comparing models 2 and 6, and 4 and 8).

The available data cannot establish whether a primary mechanism linking growth to bank structure is truly the credit underwriting practices of banks in conjunction with patterns of loan applications, as reflected in Table 3. Other factors could be at work, and the proximate effect of the lending patterns associated with higher chargeoff rates could be either positive or negative. For example, if lending occurs at more competitive interest rates in less concentrated markets, then overall investment and growth might be relatively stimulated in such markets. Greater competition might also be associated with more aggressive credit underwriting standards and hence higher chargeoff rates. However, Jayaratne and Strahan (1996) find at a statewide level that higher economic growth rates are associated with lower chargeoff rates. At any rate, Tables 3 and 5 together indicate at a minimum that the higher chargeoff rates associated with less concentrated market structures at the MSA level do not undermine local economic growth on balance, even if they constitute a net transfer of wealth from the banking sector to the real sector. Further study of this important question is warranted.

4. Conclusion

This paper has explored linkages among bank structure, loan performance, and local economic growth. Building on theoretical models of Broecker (1990) and Nakamura (1993), it has identified additional dimensions of the adverse borrower selection effects that the banking industry would be expected to exhibit as a function of structure and has provided the first empirical tests of such effects. The results indicate that newly chartered banks experience substantially higher loan chargeoff rates during their third through ninth years, consistent with the theory. Among mature banks, those operating in less concentrated banking markets experience significantly higher chargeoff rates for commercial loans and for total loans, again as predicted theoretically.

However, these higher chargeoff rates are not enough to undermine local economic growth. On the contrary, household money income was found to grow significantly faster in MSAs containing more banks, compared with MSAs containing fewer banks. Available data did not permit a precise explanation of this effect, but we cannot rule out the possibility that additional lending by banks in unconcentrated markets was a causal factor in spite of the higher average chargeoff rate. The theoretically ambiguous linkage between chargeoffs and bank profits is consistent with this interpretation.

The evidence presented here suggests at least two additional, but somewhat opposing, roles of antitrust policies in the banking industry. First, to the extent that higher chargeoff rates undermine banks' safety and soundness, structural policies that encourage lower chargeoff rates could benefit the banking industry. Such policies might include more stringent standards for chartering new banks, given the severity and duration of their loan quality problems documented above and their correspondingly reduced profitability as found by DeYoung and Hasan (1997). At the same time,

however, local economic growth is found to benefit from an unconcentrated local banking structure and could, by implication, benefit from policies that discourage local structural consolidation. Given that many of the 3,000 new U.S. banks in recent years have formed because of displaced bankers and dissatisfied customers in the wake of bank mergers, it is likely that the second policy could actually complement the first.

**Table 1: Net Chargeoff Ratio vs. Bank Age
(Quarterly Data, 1986-1995)**

<i>Variable</i>	<i>Coefficient</i>	<i>t-ratio</i>	<i>% Increment in Chargeoff Rate vs. All Banks</i>
Constant	0.00221	7.44*	--
Year 1	-0.00126	-2.68*	-60.38%
Year 2	-0.00041	-0.94	-19.57%
Year 3	0.00202	4.93*	96.89%
Year 4	0.00169	4.30*	80.96%
Year 5	0.00177	4.48*	84.78%
Year 6	0.00214	5.28*	102.15%
Year 7	0.00117	2.81*	55.84%
Year 8	0.00076	1.77**	36.60%
Year 9	0.00083	1.84**	39.90%
Year 10	0.00059	1.25	28.13%
Loans	1.570E-10	2.85*	--

*significant at the 0.01 level. **significant at the 0.10 level. Calendar time dummies not reported for brevity. Adjusted R-square = 0.0016. Number of observations = 475,027. Observations reporting negative gross chargeoffs (2494) were deleted.

**Table 2: Gross Chargeoff Ratio vs. Bank Age
(Quarterly Data, 1986-1995)**

<i>Variable</i>	<i>Coefficient</i>	<i>t-ratio</i>	<i>% Increment in Chargeoff Rate vs. All Banks</i>
Constant	0.00286	9.51*	--
Year 1	-0.00175	-3.67*	-64.09%
Year 2	-0.00089	-2.01**	-32.54%
Year 3	0.00178	4.28*	65.34%
Year 4	0.00151	3.78*	55.20%
Year 5	0.00166	4.14*	60.87%
Year 6	0.00224	5.45*	81.99%
Year 7	0.00118	2.80*	43.19%
Year 8	0.00075	1.70***	27.31%
Year 9	0.00085	1.84***	31.11%
Year 10	0.00062	1.31	22.77%
Loans	1.579E-10	2.82*	--

*significant at the 0.01 level. **significant at the 0.05 level. ***significant at the 0.10 level. Calendar time dummies not reported for brevity. Adjusted R-square = 0.0017. Number of observations = 475,027. Observations reporting negative gross chargeoffs (2494) were deleted.

Table 3: Mature Banks' Gross Chargeoff Rates and Market Structure (1990)
(t-statistics in Parentheses)

Dependent Variable: Net Chargeoff Ratio for:

<i>Loan Type:</i>	<u>Total</u>		<u>Commercial</u>		<u>Consumer</u>		<u>Other</u>	
Explanatory Variable:								
Constant	0.0038 (1.14)	-0.0008 (-0.26)	0.0032 (0.49)	-0.0024 (-0.38)	0.173 (0.75)	0.102 (0.70)	-1.50 (-0.66)	-1.42 (-0.65)
Number of banks in same MSA	1.01E-5 (2.89)*	--	1.20E-5 (1.73)**	--	(0.05)	1.12E-5 (0.35)	--	-8.20E-4 (-0.35)
log (# banks)	--	0.00149 (5.72)*	--	0.00173 (3.20)*	(0.75)	0.00936 (0.16)	--	0.0303
% high school grad. in MSA adult pop.	0.0154 (3.40)*	0.0149 (3.69)*	0.0280 (3.14)*	0.0257 (3.10)*	(-0.50)	(-0.51)	-0.153 (-0.18)	-0.098 (-0.14)
# bank employees / \$ total bank assets	3.256 (3.28)*	1.819 (2.05)**	7.308 (3.72)*	5.144 (2.82)*	(-1.25)	-114.53 (-1.22)	-70.65 (1.80) ^a	1248.2 (1.81) ^a
MSA employment / MSA population	-2.0E-7 (-4.59)*	-2.3E-7 (-5.75)*	-3.6E-7 (-4.06)*	-3.7E-7 (-4.40)*		1.2E-6 (0.40)	9.4E-8 (0.05)	3.4E-5 (1.15)
Adjusted R-square	0.0117	0.0178	0.0111	0.0106		-0.0007	-0.0005	0.0001
Number of obs.	2864		2835	2835	2814	2814	2848	2848

*significant at the 0.01 level. **significant at the 0.05 level. ^asignificant at the 0.10 level.

Sample = all banks more than five years old with assets between \$3 million and \$3 billion operating within a single MSA in 1990. The varying sample size across columns reflects the fact that not all banks reported chargeoffs in every loan category. Results for total loans are robust with respect to inclusion or omission of total MSA population, which exhibits a coefficient not significantly different from zero when included (not shown).

Table 4: Summary Statistics on MSA Growth Dataset

Variable	Mean	Std. Dev.	Min	Max
10-year growth (%)	2.57	11.49	-30.6	31.2
# banks in 1979	17.10	20.79	1	153
# banks in 1984	16.68	19.34	1	175
MSA population in 1980 (000)	593.36	1016.76	63.00	8275.00
% of MSA's adults who completed college	19.76	6.28	9.50	44.00

Table 5: Effect of Bank Structure on Local Economic Growth

Dependent variable: Percentage growth in money income per household by MSA, 1979-89.
(Heteroskedastic-consistent t-statistics in parentheses)

Variable	Model							
	1	2	3	4	5	6	7	8
Intercept	-13.16 (-5.70)*	-18.63 (-7.19)*	-13.19 (-5.75)*	-19.28 (-7.82)*	-12.06 (-5.75)*	-18.40 (-7.41)*	-12.14 (-5.85)*	-18.60 (-8.06)*
#bks79	0.126 (1.71)**	--	0.135 (3.45)*	--	--	--	--	--
log(#bks79)	--	3.411 (4.12)*	--	3.770 (5.68)*	--	--	--	--
#bks84	--	--	--	--	0.106 (1.29)	--	0.129 (2.66)*	--
log(#bks84)	--	--	--	--	--	3.816 (4.34)*	--	3.925 (5.72)*
Population	0.00018 (0.14)	0.00046 (0.57)	--	--	0.00049 (0.40)	0.00015 (0.18)	--	--
%College	0.664 (5.82)*	0.625 (5.76)*	0.665 (5.38)*	0.631 (5.84)*	0.636 (6.11)*	0.592 (6.02)*	0.636 (6.12)*	0.593 (6.04)*
Number of observations	237	237	237	237	286	286	286	286
Adjusted R-square	0.219	0.250	0.223	0.252	0.197	0.240	0.200	0.243

*significant at the 0.01 level; **significant at the 0.10 level.

References

- Altman, Edward I. and Robert Haldeman, 1995, "Corporate Credit-Scoring Models: Approaches and Tests for Successful Implementation," *Journal of Commercial Lending*, May, 10-22.
- Asch, Latimer, 1995, "How the RMA/Fair, Isaac Credit-Scoring Model Was Built," *Journal of Commercial Lending*, June, 10-16.
- Avery, Robert B., Raphael W. Bostic, Paul S. Calem, and Glenn B. Canner, 1996, "Credit Risk, Credit Scoring, and the Performance of Home Mortgages," *Federal Reserve Bulletin* 82, July, 621-648.
- Broecker, Thorsten, 1990, "Credit-Worthiness Tests and Interbank Competition," *Econometrica* 58 (2), March, 429-452.
- DeYoung, Robert and Iftekhar Hasan, 1997, "The Performance of De Novo Commercial Banks," Economics Working Paper 97-3, Comptroller of the Currency, February.
- Ferguson, Michael F. and Stephen R. Peters, 1995, "What Constitutes Evidence of Discrimination in Lending?" *Journal of Finance* 50 (2), June, 739-748.
- Hughes, Joseph P., William Lang, Loretta J. Mester, and Choon-Geol Moon, 1996, "Efficient Banking under Interstate Branching," *Journal of Money, Credit, and Banking* 28 (4), Part 2, November, 1045-1071.
- Jayaratne, Jith and Philip E. Strahan, 1996, "The Finance-Growth Nexus: Evidence from Bank Branch Deregulation," *Quarterly Journal of Economics* 111 (3), August, 639-670.
- Krol, Robert and Shirley Svorny, 1996, "The Effect of the Bank Regulatory Environment on State Economic Activity," *Regional Science and Urban Economics* 26 (5), August, 531-541.
- Mester, Loretta J., 1993, "Efficiency in the Savings and Loan Industry," *Journal of Banking and Finance* 17 (2-3), April, 267-286.
- Nakamura, Leonard I., 1993, "Loan Screening Within and Outside of Customer Relationships," Federal Reserve Bank of Philadelphia Working Paper No. 93-15, June.
- Nathan, Alli and Edwin H. Neave, 1989, "Competition and Contestability in Canada's Financial System: Empirical Results," *Canadian Journal of Economics* 22, August, 576-594.
- Shaffer, Sherrill, 1993a, "A Test of Competition in Canadian Banking," *Journal of Money, Credit, and Banking* 25 (1), February, 49-61.

Shaffer, Sherrill, 1993b, "Can Megamergers Improve Bank Efficiency?" *Journal of Banking and Finance* 17 (2-3), April, 423-236.

Shaffer, Sherrill, 1996, "Evidence of Discrimination in Lending: An Extension," *Journal of Finance* 51 (4), September, 1551-1554.

Shaffer, Sherrill and James DiSalvo, 1994, "Conduct in a Banking Duopoly," *Journal of Banking and Finance* 18 (6), December, 1063-1082.

Stiglitz, Joseph E. and Andrew Weiss, 1981, "Credit Rationing in Markets with Imperfect Competition," *American Economic Review* 71 (3), June, 393-410.

Footnotes

1. As in Ferguson and Peters (1995), this calculation incorporates the simplifying but innocuous assumptions that the cost of funds and other resource costs have been normalized to zero and that there is no partial recovery of defaulted loans.
2. Of course, this result depends in part on the maintained assumption of a fixed interest rate, whereas in practice it is possible that the degree of price competition in loans might increase with the number of banks in the market. However, several empirical studies such as Nathan and Neave (1989), Shaffer (1993a), and Shaffer and DiSalvo (1994) have found evidence of essentially competitive conduct in structurally concentrated banking markets.
3. Some other, generally dominant, motives for consolidation include cost reduction and diversification. Adverse borrower selection would not typically motivate cross-market consolidation.
4. Of course, other factors such as lower cost may also motivate the use of credit scoring models.
5. Note also that a common filter is not the same as mere knowledge that an applicant has been previously rejected by another lender, as in Nakamura's "hierarchical lenders" scenario. The latter case is consistent with i.i.d. signals for each lender, with each bank putting more weight on its own signal than on the information implied by a prior rejection.
6. Here we may relax the assumption that $r\theta_L - 1 + \theta_L < 0$ and merely assume that $r\theta_L + \theta_L < r\theta_H + \theta_H$. Note that, as documented in Shaffer (1996), a uniform interest rate r *must* be charged to both types of borrowers according to current U.S. federal policy under this interpretation of the model. Note also that, although the structure of our model resembles that of Ferguson and Peters (1995) and Shaffer (1996), the interpretation of our signals as applied to fair lending in this section is different.
7. A number of cost studies of depository institutions have excluded from the sample any institution that is less than five years old or so, on the grounds that the balance sheets and income statements of newer banks may be atypical (see for example Mester, 1993; Shaffer, 1993b; Hughes et al., 1996).
8. The year is chosen to permit contemporaneous census data to be used as control variables.
9. The adjusted R-square is lower in the logarithmic equation because each of the three demographic variables exhibits less significance than in the purely linear equation.