JOURNAL OF FINANCIAL INTERMEDIATION **7,** 359–392 (1998)
ARTICLE NO. JF980251

The Winner's Curse in Banking*

Sherrill Shaffer

Department of Economics and Finance, University of Wyoming, Laramie, Wyoming 82071-3985
E-mail: shaffer@uwyo.edu

Received January 12, 1998

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 study 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. The evidence supports the simple theory regarding loan loss rates but indicates a positive association between bank structure and income growth. *Journal of Economic Literature* Classification Numbers: G21, D80, L10. © 1998 Academic Press

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. The least risky loan applications will tend to be approved by the first bank approached. If credit screening is imperfectly correlated across banks and each lender is unaware of whether an applicant has been rejected by other banks, riskier applicants can shop around until some bank is willing to extend a loan. The more

* Much of this study was completed while the author was affiliated with the Federal Reserve Bank of Philadelphia. This article 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; for insightful comments and helpful suggestions from Mitchell Petersen, Anjan Thakor, seminar participants at the Federal Reserve Bank of Chicago, and two anonymous referees; and for helpful discussions with Scott Hoover, Fred Sterbenz, and Mark Vaughan. An earlier version of this study was presented at the American Economic Society Meetings in Chicago.

1042-9573/98 $25.00

banks are in the market, the more chances the worst applicants have of being mistaken for a good risk. The average creditworthiness of the pool of applicants is then systematically degraded as a function of the number of banks.

While the existing work on the winner's-curse problem in banking is interesting, it overlooks at least three theoretical aspects of this effect. First, to the extent that banks rely on common filters (shared databases and uniform screening criteria) in assessing applications, adverse selection can be theoretically mitigated. A corollary is that credit bureau databases and commercially standardized credit scoring models 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. This leads us to the first unresolved question I address in this paper:

- What is the impact of banks' use of common filters on the adverse selection problem in lending and what are the factors that affect banks' preferences for using common filters?

Second, the above discussion suggests that the number of banks in the market should affect the severity of the adverse selection problem, and *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 widely recognized by practitioners, but has never been formally quantified.[1] 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. This leads us to the second set of unresolved questions I address in this paper:

- How does the number of banks in the market affect an individual bank's loan loss rate? Is the loan loss rate higher for *de novo* banks?

Third, though previous studies have interpreted noisy credit screening as resulting from unobservable correlates of loan performance, 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. Since the value of these observable correlates may depend on how many banks are conditioning their decisions on them, this leads us to the third question I address:

[1] The chief financial officer of a new bank once told the author that "as soon as you open your doors, every deadbeat in town lines up to try to borrow from you" and that the only solution to this problem was to hire "superior" loan officers. Bankers and bank examiners alike are very familiar with this phenomenon.

• Are banks' profit incentives to resist particular forms of consumer legislation affected by the number of lenders in their local markets?

In addition to addressing these questions theoretically, I also develop two forms of empirical evidence of the winner's curse in banking and explore some indications of its broader macroeconomic effects. The rest of the paper is organized as follows. Section 2 develops the theory. Section 3 presents the empirical analysis. Section 4 concludes.

2. THEORETICAL CONSIDERATIONS

For expositional clarity, I present the simplest model consistent with the effects to be shown. Consider a market with N potential borrowers per period and n banks. Borrowers are of two types as in Broecker (1990) and Thakor (1996). "Good" borrowers repay a fraction θ_H of their loans on average, "bad" borrowers repay a fraction θ_L of their loans on average, and $1 > \theta_H$ ("high") $> \theta_L$ ("low") > 0. A known fraction α of these N potential borrowers are good. Banks lend at an interest rate r that is fixed, as in Ferguson and Peters (1995) and Shaffer (1996a), and exogenous, as in Nakamura (1993).

A bank, under these assumptions, 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.[2] To ensure a sustainable equilibrium, I assume that $r\theta_H - 1 + \theta_H > 0$. To rule out the trivial equilibrium in which no bank would bother to screen applicants, I assume that $r\theta_L - 1 + \theta_L < 0$.

A bank observes a noisy signal of each applicant's type and lends if and only if the signal indicates that the applicant's type is "good" (H). Signals observed by multiple banks for a given borrower are assumed to be i.i.d., where the signal corresponds to the true type with probability p_{hH} for good borrowers and p_{lL} for bad borrowers. Because an uninformative signal would have $p_{hH} = p_{lL} = \frac{1}{2}$, I assume $\frac{1}{2} < p_{hH} < 1$ and $\frac{1}{2} < p_{lL} < 1$, so that signals are noisy but 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.

2.1. *Lending with One or Two Banks*

If only one bank is in the market, all potential borrowers apply there. The bank deems $p_{hH}\alpha N + (1 - p_{lL})(1 - \alpha)N$ applicants creditworthy and

[2] 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.

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 bank's expected loan loss rate 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, I assume that half of the potential borrowers initially apply to one bank and half to the other bank. This symmetric case corresponds to Nakamura's (1993) "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 (not previously rejected by another bank) and subsequent applicants (previously rejected by another bank). Realism requires this assumption for many types of loans, especially in markets with many banks and many borrowers where it is impractical for all bankers to communicate with each other about significant fractions of applicants.[3] One likely exception occurs in consumer loans, where recent credit inquiries reported by consumer credit bureaus may indicate previous rejections. Another possible exception may involve large, highly visible borrowers such as major multinational corporations or foreign countries. Also, if a lender has a customer relationship with a borrower, then it can assume that it is the first lender approached (Nakamura, 1993). Some of our empirical results below will indirectly address these possible exceptions by disaggregating loan types for mature banks.

Under the maintained assumption, in addition to 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 applicants previously rejected by the other bank. The aggregate number of loans is $p_{hH}\alpha N(2 - p_{hH}) + (1 - p_{lL})(1 + p_{lL})(1 - \alpha)N$, 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

[3] Thus, even if bankers in a small town discuss their applicants on the golf course, this channel of communication is necessarily less complete in a large urban market. Incorporating this effect in a more generalized assumption would actually strengthen the results derived below, putting banks in an unconcentrated market at a more severe informational disadvantage than our maintained assumption would imply, compared with banks in concentrated markets.

the industry) is $[p_{\mathrm{hH}}\alpha(2 - p_{\mathrm{hH}})(1 - \theta_{\mathrm{H}}) + (1 - p_{\mathrm{lL}})(1 + p_{\mathrm{lL}})(1 - \alpha)(1 - \theta_{\mathrm{L}})]/[p_{\mathrm{hH}}\alpha(2 - p_{\mathrm{hH}}) + (1 - p_{\mathrm{lL}})(1 + p_{\mathrm{lL}})(1 - \alpha)]$.

2.2. *Lending with n Banks*

When 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

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

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

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

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

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

which is an increasing function of n since $(1 - p_{\mathrm{hH}})^{m-1} > 0$ and $p_{\mathrm{lL}}^{m-1} > 0$ for all $m \geq 1$. That is, as more banks occupy the market, fewer applicants are ultimately unable to obtain a loan. Because $p_{\mathrm{hH}} > \frac{1}{2} > 1 - p_{\mathrm{lL}}$, a higher proportion of good applicants than of bad applicants are accepted in each round, so 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 further analyzed below, implies that expected loan loss rates are an increasing function of the number of banks, as shown below:

Expected loss rate

$$= \left[p_{\mathrm{hH}}\alpha(1-\theta_{\mathrm{H}}) \sum_{m=1}^{n} (1-p_{\mathrm{hH}})^{m-1} + (1-p_{\mathrm{lL}})(1-\alpha)(1-\theta_{\mathrm{L}}) \sum_{m=1}^{n} p_{\mathrm{lL}}^{m-1} \right]$$

$$\Bigg/ \left[p_{\mathrm{hH}}\alpha \sum_{m=1}^{n} (1-p_{\mathrm{hH}})^{m-1} + (1-p_{\mathrm{lL}})(1-\alpha) \sum_{m=1}^{n} p_{\mathrm{lL}}^{m-1} \right] \quad (4)$$

$$= \{\alpha(1-\theta_{\mathrm{H}})[1-(1-p_{\mathrm{hH}})^n] + (1-\alpha)(1-\theta_{\mathrm{L}})(1-p_{\mathrm{lL}}^n)\}$$

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

The change in the expected loss rate when a market with n banks is compared to one with $n-1$ banks reduces after some algebra to

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

$$= p_{\mathrm{hH}}(1-p_{\mathrm{lL}})\alpha(1-\alpha)(\theta_{\mathrm{H}}-\theta_{\mathrm{L}}) \left[p_{\mathrm{lL}}^{n-1} \sum_{m=1}^{n-1} (1-p_{\mathrm{hH}})^{m-1} \right. \quad (5)$$

$$\left. - (1-p_{\mathrm{hH}})^{n-1} \sum_{m=1}^{n-1} p_{\mathrm{lL}}^{m-1} \right],$$

which has the sign of the expression in square brackets. We can now state the following result.

PROPOSITION 1. *An increase in the number of banks monotonically increases the expected loan loss rate.*

Proof. Examining (5) and recalling that $p_{\mathrm{lL}} > \frac{1}{2} > (1-p_{\mathrm{hH}})$, we note that $p_{\mathrm{lL}}^a(1-p_{\mathrm{hH}})^b - (1-p_{\mathrm{hH}})^a p_{\mathrm{lL}}^b = p_{\mathrm{lL}}^b(1-p_{\mathrm{hH}})^b[p_{\mathrm{lL}}^{a-b} - (1-p_{\mathrm{hH}})^{a-b}] > 0$ 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 desired result. ■

The following result is also useful to record:

PROPOSITION 2. *Despite the fact that an increase in the number of banks increases the expected loan loss rate, expected aggregate bank profits could increase as the number of banks increases.*

Proof. Expected profits for each bank i are

$$E\pi_{\mathrm{i}} = p_{\mathrm{hH}}\alpha(N/n)(r\theta_{\mathrm{H}} - 1 + \theta_{\mathrm{H}}) \sum_{m=1}^{n} (1-p_{\mathrm{hH}})^{m-1}$$

$$+ (1-p_{\mathrm{lL}})(1-\alpha)(N/n)(r\theta_{\mathrm{L}} - 1 + \theta_{\mathrm{L}}) \sum_{m=1}^{n} p_{\mathrm{lL}}^{m-1} \quad (6)$$

$$= \alpha(N/n)(r\theta_{\mathrm{H}} - 1 + \theta_{\mathrm{H}})[1-(1-p_{\mathrm{hH}})^n]$$

$$+ (1-\alpha)(N/n)(r\theta_{\mathrm{L}} - 1 + \theta_{\mathrm{L}})(1-p_{\mathrm{lL}}^n).$$

Since $1 - (1 - p_{\mathrm{hH}})^n > 1 - (1 - p_{\mathrm{hH}})^{n-1}$ for all $n > 1$, the contribution of good borrowers to expected profits increases with n. But, since $1 - p_{\mathrm{lL}}^n > 1 - p_{\mathrm{lL}}^{n-1}$ for all $n > 1$ and $(r\theta_{\mathrm{L}} - 1 + \theta_{\mathrm{L}}) < 0$ by assumption, losses to bad borrowers also rise with n. I assess how these contrary effects net out across the market by calculating the change in aggregate expected profits, comparing a market with n banks to one with $n - 1$ banks:

$$\begin{aligned}\Delta E\left(\sum \pi_{\mathrm{i}}\right) = {} & p_{\mathrm{hH}}(1 - p_{\mathrm{hH}})^{n-1}\alpha N(r\theta_{\mathrm{H}} - 1 + \theta_{\mathrm{H}}) \\ & + p_{\mathrm{lL}}^{n-1}(1 - p_{\mathrm{lL}})(1 - \alpha)N(r\theta_{\mathrm{L}} - 1 + \theta_{\mathrm{L}}).\end{aligned} \tag{7}$$

Since $1 > p_{\mathrm{hH}} > \frac{1}{2}$ and $1 > p_{\mathrm{lL}} > \frac{1}{2}$, $p_{\mathrm{hH}}(1 - p_{\mathrm{hH}})^{n-1} < p_{\mathrm{lL}}^{n-1}(1 - p_{\mathrm{lL}})$ for sufficiently large n. Also, $(r\theta_{\mathrm{H}} - 1 + \theta_{\mathrm{H}}) > 0$ and $(r\theta_{\mathrm{L}} - 1 + \theta_{\mathrm{L}}) < 0$ by assumption. Even so, the sign of Eq. (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. ■

If aggregate bank profits do increase with the number of banks in the market, then assuming that each applicant also derives utility from a loan gives us the result that social welfare is an increasing function of the number of banks, even under the maintained hypothesis of a fixed interest rate on loans—and despite the higher loan loss rates caused by adverse borrower selection with more banks. Market concentration would then cause a welfare loss for reasons unrelated to pricing; any market power associated with concentration would exacerbate this welfare loss.

2.3. *Discussion of Key Assumptions*

Two of the most restrictive assumptions in the partial equilibrium model above include a fixed interest rate on loans and binary signals. The assumption that banks are ignorant of prior rejections also warrants discussion.

The loan rate assumption is maintained for tractability in obtaining sharp results, for expositional transparency, for consistency with some related theoretical literature (Ferguson and Peters, 1995; Shaffer, 1996a; Nakamura, 1993), and to some extent for realism. Virtually all structural models or other "new empirical industrial organization" studies of banking find essentially competitive loan pricing—not only nationwide (Shaffer, 1989, 1996b) but also in highly concentrated markets such as Canada where the industry is dominated by fewer than half a dozen large banks (Nathan and Neave, 1989; Neave and Nathan, 1991; Shaffer, 1993a), rural counties in a unit banking state (Shaffer, 1985), or even a banking duopoly (Shaffer and DiSalvo, 1994).[4]

Even so, it is useful to discuss the possible effect of endogenizing the loan rates. In any symmetric equilibrium with binary signals, the proportion of defaulting loans will not be a function of the interest rate on loans except to the extent that a borrower's propensity to default is itself a function of the interest rate. Even then, if a bank could reduce its default rate by lowering its interest rate, it would have an incentive to do so in both concentrated and unconcentrated markets. Moreover, some studies such as Hannan and Liang (1993) have found evidence that banks tend to pay higher interest rates on deposits in less concentrated markets, thereby incurring higher costs that would reduce the potential for charging lower interest rates on loans in such markets (though see the previous footnote for doubts about this result). Similarly, the higher loan loss rates shown above constitute another element of cost that would be higher in less concentrated markets. Thus, even if concentrated markets permit some market power in loan pricing, the sign of the relationship between equilibrium loan rates and the number of banks could depend on parameter values and so would be an empirical question. With regard to loan loss rates, the empirical section of this article below finds that the data support the predictions of the simple partial equilibrium model.

The assumption of binary signals representing outcomes of a bank's loan screening process—as in Thakor (1996)—is both parsimonious and realistic, in that the outcome of any bank's credit screening process always reduces to a binary decision to lend or not to lend. What is obscured in the binary

[4] Some of these studies use bank-specific data while others use industry aggregate data. Structural models yield valid results for either type of data and—provided that the sample spans a sufficiently broad market or set of markets—regardless of the size of the true geographic market; see Bresnahan (1982), Lau (1982), and Shaffer (1993a, footnote 7). Likewise, some such studies have used data from individual local geographic banking markets, including Shaffer and DiSalvo (1994) and Shaffer (1985). If conduct is competitive and if costs do not vary systematically with market structure, then the interest rate on loans will not be a function of the number of banks. On the other hand, Berger and Hannan (1989) and Hannan and Liang (1993) find empirical evidence of market power in the pricing of bank deposits, the former using a traditional structure–conduct–performance (SCP) approach and the latter using a "new empirical industrial organization" model based on Panzar and Rosse (1987). However, such a finding need not suggest market power on the loan side for several reasons. First, many depositors are more locally limited and face fewer alternatives than borrowers. Second, the SCP approach is well known to admit alternative interpretations, and the Panzar–Rosse test has been shown to be a one-sided test that can spuriously indicate market power if the sample is not in long-run equilibrium (Shaffer, 1983; Shaffer and DiSalvo, 1994). Third, the structural models that found essentially competitive pricing in Shaffer (1989, 1996b) and Shaffer and DiSalvo (1994) have been shown to reflect any market power on *either* the input or output sides (Shaffer, 1999). These last two considerations cast doubt on Hannan and Liang's findings. Berger (1991) examines the price–concentration relationship that remains in banking after controlling for efficiency, and concludes that the structure–conduct performance paradigm "may have some validity in deposit markets, but not in loan markets" (page 25).

stylization is a bank's ability to adjust its lending criteria in response to perceived risk factors, which could include the number of rival lenders. No formal studies have addressed whether bankers in practice consider market structure in setting their lending standards. However, anecdotal evidence from bankers suggests that they do respond to the *practices*, but not consciously to the *number*, of rival lenders. Thus, although in principle bankers could adopt more conservative lending standards in markets with larger numbers of banks—thereby partially mitigating the adverse selection effects shown below—in practice it appears that the model here may provide a more realistic characterization of lending behavior. A similar conclusion applies to *de novo* banks, except that they must balance the need for caution against the need to expand their loan portfolio rapidly in order to realize economies of scale and diversification.[5]

Broecker (1990) likewise assumes binary screening outcomes. Nakamura (1993) endogenizes the acceptance threshold in a continuous (normal) density function of signals, but does not explore the impact of market structure on loan loss rates per se. In a related analysis of refereed journals' screening of articles, Nakamura and Shaffer (1993) have shown the intuitive result that endogenous acceptance thresholds only partially offset the adverse selection problem related to structure. Thus, abstracting from endogenous acceptance thresholds will overstate the severity of the adverse selection problem, but will not alter its sign. Here too, the issue is empirical.[6]

[5] Although the minimum efficient scale of a commercial bank remains controversial, the consensus from the empirical literature is that there are significant economies of scale up to at least \$100 million of assets (Berger and Humphrey, 1992), and some studies provide much larger estimates of that scale (Hughes and Mester, 1998).

[6] A referee has noted that, even within the framework of binary signals, a bank can alter its effective acceptance rate by choosing to apply multiple binary tests per applicant—and could even have an incentive to do so in cases where the sign of Eq. (7) is positive, since a given number of banks (even a monopoly bank) could thereby reap the profit benefits otherwise associated with larger numbers of banks. Something similar to this seems to occur where banks apply a multilayer internal loan review, though many banks use that process only to increase the number of approvals and not to support rejections. However, such a process would differ from that involving multiple banks in two key ways. First, each bank knows the results of all of its own layers of review, not only in practice but also in a theoretical optimum (not modeled here) since self-imposed ignorance of other internal reviews would involve discarding useful information. Second, the outcomes of multiple internal reviews will likely be more highly correlated than the outcomes of reviews across banks, since any substantial variation across internal reviews would seem to require that one or more of those reviews must ignore relevant information available within the bank. The appropriate model of multiple internal reviews would then require knowledge of prior internal outcomes (a case that has been analyzed, though not in great detail, by Nakamura, 1993) but also could not assume i.i.d. conditional distributions of signals (a case that has not been formally analyzed in extant literature). In addition, multiple internal reviews with known prior outcomes would be subject to the "informational cascade" phenomenon (Bikhchandani *et al.*, 1992): beyond some point, subsequent levels of review would yield outcomes that are marginally uninformative, given the information contained in the aggregation of prior reviews and would thus contribute nothing to the screening process.

Recall that the calculations above assume that each bank is ignorant of an applicant's prior rejections. Intuitively, we should expect that the adverse selection of borrowers would be mitigated if banks are able to aggregate information across multiple applications. Indeed, one obtains quantitatively different results if banks are aware of multiple applications even without being able to infer other banks' signals, as could occur when borrowers apply to several lenders simultaneously; see Thakor (1996). However, related analysis by Nakamura and Shaffer (1993) suggests that even when a bank's awareness of prior applications is coupled with the additional knowledge of prior rejections, the resulting aggregation of information may not always suffice to offset the adverse selection completely. Again, the extent to which adverse selection is mitigated, whether by information aggregation or by other mechanisms, is an empirical issue.

2.4. *Applicant Attrition*

A further detail of adverse borrower selection 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 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 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. Long delays in granting some business loans may impair the borrower's financial condition, further undermining her ability to qualify for a loan. A succession of rejections may also lead an applicant to expect future rejections. For all these reasons, any 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 financial information available for 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 enough times, the likelihood of eventual acceptance approaches zero. Such nonlinearity is empirically tested below using a concave function (the natural logarithm) of the number of banks.

2.5. *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, Experian/TRW, or Trans Union). Thus, every bank 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 credit bureaus' databases. 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.[7]

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,

[7] Some other, generally dominant, motives for consolidation include cost reduction and diversification. Adverse borrower selection would not typically motivate cross-market consolidation.

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 by mitigating adverse selection. The more banks operate in a market, the more this benefit of a common filter 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.[8]

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 Eq. (3) above are given by

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

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

The magnitude of this effect can be quite striking. If $p_{hH} = p_{lL} = 0.9$ and $n = 20$, then (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 are 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

[8] Of course, other factors such as lower cost and quicker turnaround may also motivate the use of credit scoring models.

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

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 Eq. (9) equals 0.000942, so 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.

Now consider the same parameter values except $\theta_{\text{L}} = 0.8$ and $\alpha = 0.9$. The right-hand side of Eq. (9) is then 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.

Another important question is whether banks would choose a common filter rather than i.i.d. filters in a noncooperative equilibrium. Specifically, if $n - 1$ banks choose a common filter with selectivities β_{hH} and β_{lL}, will the last bank also choose the common filter or instead deviate to an idiosyncratic filter with selectivities p_{hH} and p_{lL}, uncorrelated with the common filter? The exact effect of deviation on the last firm's expected profits depends on what applicants know and how they behave. The first of these issues is whether each bank's choice of common versus i.i.d. filter is common knowledge. If so, then no borrower would ever apply at more than one of the banks using a common filter; if not, then an applicant rejected by any bank would be equally likely to apply at any other bank, regardless of its filter type. In the following analysis I adopt the former assumption, though the analysis can also be carried out under the alternative.

Given this informational assumption, the second issue is whether each applicant is equally likely to apply first at a bank with a common filter or at one with an i.i.d. filter. The simplest case—which I maintain below—is the symmetric case, which might result if applicants are uncertain about the relative magnitudes of β_{hH}, p_{hH}, β_{lL}, and p_{lL}.[9] Under these conditions, the bank that chooses an i.i.d. filter (given that $n - 1$ other banks choose a common filter) expects to earn

$$E\pi_d = p_{hH}\alpha N[1 + (1 - \beta_{hH})(n - 1)](r\theta_H - 1 + \theta_H)/n + (1 - p_{lL})(1 - \alpha)N[1 + \beta_{lL}(n - 1)](r\theta_L - 1 + \theta_L)/n. \quad (10)$$

The bank will deviate from the common filter if and only if this expression exceeds the expected profit using a common filter, equal to

$$E\pi_c = \beta_{hH}\alpha N(r\theta_H - 1 + \theta_H)/n + (1 - \beta_{lL})(1 - \alpha)N(r\theta_L - 1 + \theta_L)/n. \quad (11)$$

The incremental expected profit from unilateral deviation is $D \equiv E\pi_d - E\pi_c$, or

$$D = \alpha N(r\theta_H - 1 + \theta_H)\{p_{hH}[1 + (1 - \beta_{hH})(n - 1)] - \beta_{hH}\}/n + (1 - \alpha)N(r\theta_L - 1 + \theta_L)\}(1 - p_{lL})[1 + \beta_{lL}(n - 1)] - (1 - \beta_{lL})\}/n. \quad (12)$$

D is an increasing function of p_{hH} and p_{lL} but a decreasing function of β_{hH} and β_{lL}. That is, as intuition suggests, a bank is more likely to choose an i.i.d. filter to the extent that it is intrinsically more accurate than the common filter. The sign of D depends on all parameter values in a way that is not straightforward to interpret. Thus, in general, the choice of a common filter can be either a noncooperative equilibrium or not, according to the relative selectivities of the two filters in conjunction with the relative mix and profitability of the two types of borrowers.[10]

The rational outcome may also vary systematically across types of bank loans, depending on the relevant geographic market for each. For instance,

[9] Generally, as explored below, either $p_{hH} > \beta_{hH}$ or $p_{lL} > \beta_{lL}$ is required to prompt a bank's deviation, but not necessarily both. The alternative to the symmetric case occurs if borrowers believe that both of the above conditions hold; then type H borrowers will apply first to the i.i.d bank in order to maximize their likelihood of receiving a loan on the first try, whereas type L applicants will choose to apply first to a bank using a common filter in order to maximize their likelihood of being mistaken for a good borrower on the first round.

[10] In the instructive special case where $p_{hH} = \beta_{hH} = p_{lL} = \beta_{lL}$, D takes the sign of $\alpha(r\theta_H - 1 + \theta_H) + (1 - \alpha)(r\theta_L - 1 + \theta_L)$, which is negative whenever there is any incentive for banks to screen at all. Thus, for deviation from the common filter to be profitable, the i.i.d. filter must exhibit some superiority to the common filter in terms of its intrinsic selectivity. Clearly, if it is not profitable for only one bank to deviate, then deviation by multiple banks will also be unprofitable (for a given selectivity of the i.i.d. filter), since multiple deviations compound the adverse borrower selection. If deviation is profitable for a single bank, further analysis could explore the maximum number of banks that could profitably deviate. If that number is less than n, the equilibrium selection process must also identify *which* banks would deviate, starting from an ex ante identical set of banks.

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 to lend to riskier applicants, and this learning process might initially entail some shopping around. 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.[11] Given a stable set of policies across banks, one expects that borrowers would eventually self-select in a sequential or hierarchical fashion, with the least creditworthy applying to those banks setting the lowest acceptance threshold. Such a self-selection equilibrium, which is beyond the scope of this article to analyze formally, could 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.

2.6. *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 initial and subsequent applicants common to all mature banks. The fraction of the *de novo* bank's applicant pool reflecting this backlog will depend on how long an applicant is willing to keep applying after the original rejection. While that fraction may vary, I can at least characterize the mix of types within the backlog and compare that mix with the proportions of types facing mature banks in the same market. This is done in the result below.

[11] 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 (1993) "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.

PROPOSITION 3. *If a* de novo *bank's signal has the same probability distribution as the signal of an incumbent bank, the* de novo *bank will suffer a higher loan loss rate than incumbent banks.*

Proof. The ratio of good to bad applicants within the pool of those previously rejected n times is $\alpha(1 - p_{\mathrm{hH}})^n/[(1 - \alpha)p_{\mathrm{lL}}^n]$. The corresponding ratio for those who have been rejected $n - 1$ times is $\alpha(1 - p_{\mathrm{hH}})^{n-1}/[(1 - \alpha)p_{\mathrm{lL}}^{n-1}]$. For any n, the ratio of these two ratios is $(1 - p_{\mathrm{hH}})/p_{\mathrm{lL}}$, which is less than 1 since $p_{\mathrm{hH}} > \frac{1}{2}$ and $p_{\mathrm{lL}} > \frac{1}{2}$. Thus, at each successive stage of rejection, the remaining pool of potential borrowers going forward to the next round of applications is monotonically worse. For any value of n, the backlog facing a new bank thus has a worse mix and will generate higher loan loss rates than the group of potential borrowers still applying among the incumbent banks. ∎

This proposition says that as long as the accuracy of the new bank's loan screening is similar to that of the incumbent banks, the *de novo* bank will suffer a higher loan loss rate than the mature incumbents. As discussed above, such a bank may be able to offset this problem through more conservative lending standards, but typically only partially. This prediction will be tested in the empirical section below.

2.7. *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 in aggregate with average 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, sex, 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. Two possible linkages emerge, depending on how lenders interpret the prohibited criteria.

In one case, suppose that lenders might attempt to use prohibited but readily observable criteria (such as gender or race) as the signal itself. Such a signal would be uniformly observed by every lender and thus could function as a common filter.[12] Since the results above establish that the use of a common filter can reduce loan loss rates even if the filter is intrinsically poor (i.e., only weakly correlated with actual default rates), the model here

[12] I am grateful to an anonymous referee for this observation. Banks in the U.S. cannot legally respond to such information by pricing the risk directly, because the Equal Credit Opportunity Act and the Fair Housing Act prohibit charging different interest rates for a given loan to borrowers from different groups (Interagency, 1994, page 3).

suggests an economic reason why lenders might tend to use such criteria, especially in unconcentrated markets. This scenario would imply that lenders might be more resistant to complying with fair lending laws in large urban communities containing many banks than in smaller communities—and that, in the absence of fair lending legislation, discriminatory lending patterns would be more pronounced in urban markets.[13]

If such a pattern exists, two policy conclusions follow. First, fair lending laws can succeed in deterring banks from using prohibited criteria as a common filter because, once several banks in the community are induced to stop using those criteria, such criteria become less effective as a filter for the remaining banks. Second, introducing alternative common filters such as standardized credit scoring models can reduce the economic incentive for lenders to rely on filters that have a dubious objective basis and are socially objectionable.[14]

In an alternative case, prohibited criteria might be used as a subset of inputs to the formation of a signal rather than as the signal itself. In this case, different lenders could draw different conclusions from the same observable criteria, which would no longer serve as a common filter. Then, if the prohibited criteria have any incremental predictive value (as some studies such as Berkovec *et al.*, 1996, have suggested), ignoring those criteria as required by law would reduce the accuracy of each lender's credit screening process. However, because any screening process that is imperfectly correlated across banks is less effective in less concentrated markets, the value to lenders of a given incremental improvement to the accuracy of that screen is also lower in such markets. In this case, banks in large urban markets will have less financial incentive than banks in more concentrated markets to oppose or circumvent fair lending laws.[15]

3. EMPIRICAL ANALYSIS

As suggested above, some mechanisms can potentially mitigate adverse borrower selection. Besides common filters, relationship banking is a traditional banking strategy that can reduce the problem by discouraging bor-

[13] I have been unable to document whether this pattern has existed, but conversations with federal bank examiners and other banking economists familiar with exam data have uncovered a small amount of anecdotal evidence that this pattern of resistance was seen during the early years of the fair lending laws.

[14] This point provides theoretical support for a conjecture of Galster (1996), though Ladd (1998) cautions that credit scoring models themselves might in some cases fail to guard against disparate impact on minority borrowers.

[15] 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 (1996a), 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 (1996a), the interpretation of our signals as applied to fair lending in this section is different.

rowers from shopping around. Collateral can help both by deterring bad borrowers from applying (knowing they will lose some of their collateral upon default) and by enhancing the lender's recovery in the event of default. Soliciting loans via preapproved applications can enhance a bank's chances of being the first to screen a given applicant, though this advantage fails if all banks solicit simultaneously. To the extent that banks can charge high application fees, they can mitigate adverse selection by fostering applicant attrition. Because of such possibilities, the true extent of adverse borrower selection is an empirical question.

This section tests some of the properties predicted above, plus others developed below. I explore loan chargeoff ratios first as a function of a bank's age, then as a function of the number of banks per market. Finally, I explore empirical linkages with economic growth rates.

3.1. *Age Effects*

Here I examine loan chargeoff ratios as a function of a bank's age, using data from the quarterly Call Reports filed by all banks. The sample spans all U.S. commercial banks during 1986–95 and contains nearly half a million observations. Bankers, regulators, and academic researchers have consistently considered that banks exhibit atypical performance during their first five years, so I am 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.[16] I include annual age dummies for each of a bank's first 10 years—longer than the period tracked by most previous studies of *de novo* bank performance and longer than observed periods of anomalous profitability (DeYoung and Hasan, 1998). I also include quarterly calendar time dummies to control for business cycle and other macroeconomic or environmental effects.[17]

Consistent with bankers' common perception, net chargeoff rates—as shown in Table I—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

[16] 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).

[17] While this model incorporates fixed effects with respect to time, it incorporates random effects with respect to individual banks. Maddala (1987) has shown that, when the sample is not identical with the population of interest, inferences about true population values are more efficiently carried out by random-effects models than by fixed-effects models (see also Emmons, 1993, page 193). Since banks that were new within our sample period are a proper subset of the population of interest, we estimate random-effects models only. This approach also follows DeYoung and Hasan (1998), who study *de novo* banks' performance without fixed effects.

TABLE I
Net Chargeoff Ratio vs Bank Age
(Quarterly Data, 1986–1995)

Variable	Coefficient	t-ratio	% Increment in chargeoff rate vs all banks
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	1.570×10^{-10}	2.85*	—

Note. Calendar time dummies not reported for brevity. Adjusted R-squared = 0.0016. Number of observations = 475,027 drawn from all U.S. banks over 40 quarters. Observations reporting negative gross chargeoffs (2494) were deleted.

Source. Quarterly Call Reports.

* Significant at the 0.01 level.

** Significant at the 0.10 level.

7 through 9, the chargeoff rate is still about 40 to 50% higher than that of mature banks, with the difference statistically significant. Only in a bank's 10th 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.

One might suspect that the normal seasoning of a new loan portfolio could perhaps account for this pattern of chargeoffs, quite apart from any adverse selection. However, data on loan seasoning from mature banks are not consistent with this alternative interpretation. Using data on commercial loans from the Federal Reserve Board of Governors' Quarterly Survey of Bank Lending, Avery and Gordy (1995) find that chargeoff rates peak and return to normal levels much sooner than the pattern found for *de novo* banks here.[18] The much longer and more severe increase in chargeoff rates

[18] In particular, the chargeoff rates on a representative loan portfolio subjected to rapid growth—a doubling of loan flow rates in each of four quarters leading to a 40% increase in the stock of outstanding loans by the fourth quarter—was found to yield lower chargeoff rates during the first year (similar to our *de novo* findings), a subsequent increase in chargeoff rates during the second year sufficient to restore the overall chargeoff rate to near its original level, and no further significant changes in the chargeoff rate during quarters 8 through 40. (See their Fig. 4 and conclusion on pages 2f. that "most of the impact is gone within a year after the impulse is stopped.")

TABLE II
Gross Chargeoff Ratio vs Bank Age
(Quarterly Data, 1986–1995)

Variable	Coefficient	t-ratio	% Increment in chargeoff rate vs all banks
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	1.579×10^{-10}	2.82*	—

Note. Calendar time dummies not reported for brevity. Adjusted R-squared = 0.0017. Number of observations = 475,027 drawn from all U.S. banks over 40 quarters. Observations reporting negative gross chargeoffs (2494) were deleted.

Source. Quarterly Call Reports.

* Significant at the 0.01 level.

** Significant at the 0.05 level.

*** Significant at the 0.10 level.

observed here for *de novo* banks demands a different explanation. Nor is it likely that our findings reflect merely a learning process involving inexperienced lenders: *de novo* banks always begin operation with experienced staff and management in selected key positions, as a regulatory prerequisite to chartering the bank. Thus, absent additional compelling hypotheses, the theoretically established adverse selection appears the most likely explanation of the observed phenomenon. At a minimum, 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 II shows that new 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 I. Comparing the right-hand columns of Tables I and II, 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 support and complement results on the profit efficiency of *de novo* banks. DeYoung and Hasan (1998) found that new banks are significantly less profitable than older banks of the same asset size until after the ninth year of a bank's operation. However, the pattern of loan delinquencies observed in that study did not appear to contribute to the lower profitability, and loan chargeoffs were not explored. Taken together, their study and this one suggest that the theoretically ambiguous impact of adverse borrower selection on bank profitability tends to operate in the same direction as chargeoff rates for *de novo* banks.

3.2. *Structure Effects*

A second test 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, I find that loan chargeoff rates are a significantly increasing function of the total number of banks in the MSA, consistent with the theory (Table III).[19] 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 (b.p.), 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 III also shows that the log of the number of banks is even more significant and yields a higher adjusted R-squared than the raw number of banks, implying a nonlinear association between bank structure and performance consistent with the theory of the previous section. The net chargeoff ratio rises by 2.7 b.p. as the number of banks in an MSA increases from five to six, or by 1.6 b.p. as the number of banks rises from nine to 10, for example. Thus, for MSAs with few banks, structure has a greater impact 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.

[19] The year is chosen to permit contemporaneous census data to be used as control variables. Call Reports provided bank-specific data on chargeoffs, number of employees, and bank assets; the FDIC's Summary of Deposits database was used to identify banks operating within a single MSA and to quantify the number of banks and total bank deposits in each MSA; demographic data by MSA were taken from the *1996 County and City Extra: Annual Metro, City and County Data Book* (Bernan Press, Lanham, Maryland). The relationship shown in Table III persists across total loans even when the population of the MSA (not shown) and other characteristics are controlled for.

TABLE III
Mature Banks' Gross Chargeoff Rates and Market Structure (1990)
(t-Statistics in Parentheses)

	Dependent variable: Net chargeoff ratio by loan type:							
	Total loans		Commercial		Consumer		Other	
Constant	0.0038 (1.14)	−.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	1.01×10^{-5} (2.89)*	—	1.20×10^{-5} (1.73)**	—	1.12×10^{-5} (0.05)	—	-8.20×10^{-4} (−0.35)	—
log (# banks)	—	0.00149 (5.72)*	—	0.00173 (3.20)*	—	0.00936 (0.75)	—	0.0303 (0.16)
% high school grad. in MSA adult pop.	0.0154 (3.40)*	0.0149 (3.69)*	0.0280 (3.14)*	0.0257 (3.10)*	−0.153 (−0.50)	−0.098 (−0.51)	−0.535 (−0.18)	−0.413 (−0.14)
# bank employees/$ total bank assets	3.256 (3.28)*	1.819 (2.05)**	7.308 (3.72)*	5.144 (2.82)*	−114.53 (−1.25)	−70.65 (−1.22)	1248.2 (1.80)[a]	1214.8 (1.81)[a]
MSA employment/ MSA population	-2.0×10^{-7} (−4.59)*	-2.3×10^{-7} (−5.75)*	-3.6×10^{-7} (−4.06)*	-3.7×10^{-7} (−4.40)*	1.2×10^{-6} (0.40)	9.4×10^{-8} (0.05)	3.4×10^{-5} (1.15)	2.6×10^{-5} (0.88)
Adjusted R-squared	0.0117	0.0178	0.0111	0.0106	−0.0007	−0.0005	0.0001	0.0001
Number of obs.	2864	2864	2835	2835	2814	2814	2848	2848

Note. 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).

Source. Call Reports for chargeoffs, banks assets, and bank employees; FDIC's Summary of Deposits for number of banks in each MSA; *1996 County and City Extra: Annual Metro, City and County Data Book* (Bernan Press, Lanham, MD) for demographic data.

* Significant at the 0.01 level.

** Significant at the 0.05 level.

[a] Significant at the 0.10 level.

However, the effect is not uniform across all types of loans. When chargeoffs are disaggregated by loan type (last six columns of Table III), bank structure exhibits a significant linkage with commercial loans but not with consumer loans or "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 or evidence of prior rejections (for example, from credit bureaus) for consumer loans, standardized lending criteria for some categories of real estate loans, and other factors. In addition, geographic markets for some loan types—such as credit card loans—may not be well approximated by MSAs, rendering our structural measures inaccurate in those cases. Data for savings and loan associations were unavailable and, though the number of S&Ls should be positively correlated with the number of banks in an MSA, the exclusion of S&Ls might 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.[20]

3.3. *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, is limited by law as explained in footnote 12, would carry its own adverse selection effects as analyzed by Broecker (1990), and has other limitations analyzed by Stiglitz and Weiss (1981). Thus, within a market, good borrowers must subsidize bad borrowers. The ambiguous effect of adverse selection on banks' expected profits, shown theoretically above, indicates that the net welfare effects are not obvious.

The broader effect of adverse borrower selection on the economy hinges on whether the primary distinction between good and bad borrowers involves the mean or dispersion of the returns on their respective projects. The standard loan contract limits the extent to which banks can share in the upper tail of volatile returns. If "bad projects" are more uncertain than "good projects" but have comparable mean returns, funding bad projects may benefit society on average but redistribute expected profits from the banking sector to the real sector. The recently recognized importance of the financial intermediation sector in fostering economic growth—both within the United States (Jayaratne and Strahan, 1996; Krol and Svorny, 1996) and in other countries (King and Levine, 1993; Levine, 1998; Rajan and Zingales, 1998)—underscores the relevance of this question.

This section presents regressions addressing this issue, using data summarized in Table IV. The dependent variable is the percentage change in money income per household between 1979 and 1989 (in constant dollars) by MSA for a nationwide sample, analogous to the statewide economic growth variables explored in Jayaratne and Strahan (1996) and Krol and Svorny (1996). The use of a 10-year cumulative growth rate—a single observation for each MSA—provides the advantages of smoothing out high-frequency intertemporal noise and mitigating the impact of outlier years in growth rates, as in Levine (1998) and others. Regressors include the MSA population in 1980, the percentage of adults in each MSA with four or more years of college education as of 1990 (a proxy for available human capital), and the number of banks present 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 b.p. per additional bank in the MSA. When population is included in the regression, its collinearity with the number of banks ($\rho = 0.89$ for

[20] The adjusted R-squared is lower in the logarithmic equation because each of the three demographic variables exhibits less significance than in the purely linear equation.

TABLE IV
Summary Statistics on MSA Growth Dataset

Variable	Mean	Std. Dev.	Min	Max
MSA growth rate (%) of median household money income, 1979–1989 (constant dollars)	2.57	11.49	−30.6	31.2
Number of banks in 1979	17.10	20.79	1	153
Number of banks in 1984	16.68	19.34	1	175
MSA population in 1980 (000)	593.36	1016.76	63.00	8275.00
Percentage of MSA's adults who completed college by 1990	19.76	6.28	9.50	44.00
Net entry rate by banks into MSA (ratio of number banks in 1984 to number banks in 1979)	1.145	0.331	0.200	2.286
Total bank deposits in MSA ($ millions)	4429	1439	1.29	185119
Bank deposits per capita ($000)	5.386	2.646	0.018	22.371
Growth rate (%) of state's personal income per capita, 1980–1990 (constant dollars)	14.83	7.64	−5.30	29.90

Sources. 1996 County and City Extra: Annual Metro, City and County Data Book (Bernan Press, Lanham, MD); FDIC's *Summary of Deposits; Statistical Abstract of the United States, 1995* (U.S. Bureau of the Census, Washington, D.C.).

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. Table V presents the regression results.

The even-numbered models in Table V use the natural logarithm of the number of banks in each MSA. This form is intended to capture the effect of applicant attrition as discussed above, assuming the linkage between structure and growth reflects to some extent the adverse selection of applicants. These log variables all yield larger t-ratios and adjusted R-squareds 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 income growth. The estimated coefficients imply that an additional bank is associated with an increase of 62 to 72 b.p. in the 10-year growth rate of an MSA with five banks initially, 36 to 41 b.p. for an MSA with nine banks, and 14 to 16 b.p. for an MSA with 24 banks, for example.

Because of collinearity between the number of banks and population, I

TABLE V
Association between Bank Structure and Local Economic Growth

	Model							
Variable	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-squared	0.219	0.250	0.223	0.252	0.197	0.240	0.200	0.243

Note. Dependent variable: Percentage growth in median money income per household by MSA, 1979–1989. (Heteroskedastic-consistent t-statistics in parentheses.)

Sources. 1996 County and City Extra: Annual Metro, City and County Data Book (Bernan Press, Lanham, Maryland); FDIC's *Summary of Deposits; Statistical Abstract of the United States, 1995* (U.S. Bureau of the Census, Washington, D.C.).

* Significant at the 0.01 level.

** Significant at the 0.10 level.

explored the robustness of the estimates across variants of the model.[21] Table VI reports regressions controlling for the absolute and relative size of the banking sector, as in King and Levine (1993). Aggregate scale is measured by total bank deposits in each MSA as of June 1984 (the midpoint of the sample period).[22] If the dominant factor associated with income growth rates is the absolute size of the banking sector rather than the number of institutions, then including total deposits as a regressor should

[21] Besides the variants reported in Table IV, other variants used subsets of MSAs with populations between 100,000 and 1 million, or between 100,000 and 500,000, to reduce collinearity. 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 were both greater for all 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.

[22] Deposits, measured in millions of dollars, are taken from the FDIC's Summary of Deposits.

TABLE VI
Association between Bank Structure and Local Economic Growth, Controlling for Absolute and Relative Size of the Banking Sector

	Model					
Variable	1	2	3	4	5	6
Intercept	−18.81 (−7.27)*	−18.80 (−7.63)*	−18.20 (7.32)*	−17.77 (7.57)*	−19.28 (−7.45)*	−18.44 (−7.71)*
log(#bks79)	3.64 (4.30)*	—	3.26 (4.70)*	—	2.96 (3.64)*	—
log(#bks84)	—	4.15 (4.69)*	—	3.50 (4.81)*	—	3.22 (3.78)*
Population	−0.000928 (−0.69)	−0.00170 (−1.16)	—	—	0.000341 (0.46)	0.000232 (0.32)
% College	0.626 (5.77)*	0.591 (6.02)*	0.621 (5.75)*	0.587 (6.00)*	0.595 (5.47)*	0.561 (5.72)*
Total deposits (1984, $mill.)	0.000120 (1.26)	0.000163 (1.27)	7.17×10^{-5} (3.13)*	7.08×10^{-5} (2.83)*	—	—
Deposits per capita ($000)	—	—	—	—	0.455 (1.84)***	0.369 (4.13)*
Number of observations	237	285	237	285	237	285
Adjusted R-squared	0.253	0.250	0.255	0.247	0.255	0.254

Note. Dependent variable: Percentage growth in median money income per household by MSA, 1979–1989. (Heteroskedastic-consistent t-statistics in parentheses.)

Sources. 1996 County and City Extra: Annual Metro, City and County Data Book (Bernan Press, Lanham, Maryland); FDIC's *Summary of Deposits; Statistical Abstract of the United States, 1995* (U.S. Bureau of the Census, Washington, D.C.).

* Significant at the 0.01 level.
** Significant at the 0.05 level.
*** Significant at the 0.10 level.

render the number of banks insignificant. Instead, the coefficient on total deposits is not significant while that on the log number of banks remains positive and significant at the 0.01 level.

Because total deposits are highly correlated with population ($\rho = 0.83$), additional regressions were run omitting population, as reported in the next two columns of Table VI. The coefficient on total deposits is positive and significant at the 0.01 level but again fails to diminish the significance of the log number of banks. The last two models in Table VI control for aggregate bank deposits per capita in each MSA, a measure of the relative scale of the banking sector. The coefficient on this variable is significantly positive in both models, but the coefficient on the log number of banks also remains positive and significant at the 0.01 level. These outcomes indicate that the number of banks in a community is robustly associated with income growth rates, even controlling for the absolute or relative scale of the banking sector.[23]

One can question whether these results reflect reverse causality, if bank structure responds to either actual or anticipated economic growth within a market. Growing markets should attract entry while shrinking markets could force exit. Several points are relevant here. First, income growth *per household* need not indicate overall market growth or an attraction to entry in MSAs with declining population. Second, if banks respond to actual but not anticipated growth, then the 1979 bank structure could only be endogenously related to 1979–89 economic growth if growth rates exhibit intertemporal persistence. Economic growth has been found not to exhibit intertemporal persistence in various countries (see for example Easterly *et al.* 1993) or in states across the U.S.;[24] this pattern also appears to characterize MSAs, since the correlation of growth rates of real per capita personal income from 1984–1990 and 1990–1994 is only 0.13 for the 44 largest metropolitan areas as of 1984.[25] Thus, reverse causality here would seem to require that banks adjust their structure in anticipation of future economic growth trends.

In any case, the linkage between entry or exit and market growth can be tested directly. Regressions reported in Table VII add a measure of net banking entry, the number of banks in each MSA in 1984 divided by the number in 1979. If the number of banks responds to *either* actual or anticipated income growth rates, this variable should exhibit a positive coefficient; such a finding would suggest that the linkage between bank structure and economic growth evident in the other models in Tables V and VI might be biased by reverse causality. In fact, the coefficient is significantly *negative*, indicating that bank entry tended to occur more in the MSAs experiencing slow or negative income growth than in the high-growth MSAs. At the same time, the magnitude and significance of the other coefficients (particularly on the log number of banks) is not materially affected. This finding suggests that the observed linkage between income growth rates and the number of banks is not likely an artifact of reverse causality.

[23] Other combinations of the variables shown in Table VI were tried in additional regressions not reported in the table. In general, the inclusion of any larger subset of these variables in a single regression rendered the coefficients insignificant on total deposits, per capita deposits, and bank entry but without materially altering the magnitude or significance of the coefficients on the other variables (particularly on the number of banks).

[24] The correlation between states' median family income growth over the consecutive decades 1959–1969 and 1969–1979 is only 0.212, and the earlier growth rate explains only 4.67% of the variation in the subsequent growth rate (adjusted R-squared = 0.027) in a simple regression using data from the *Statistical Abstract of the U.S.*, various years.

[25] In a simple regression, the earlier growth rate explained only 1.7% of the variation in the later growth rate, and the adjusted R-squared was negative. We should expect that a comparison across 10-year periods would exhibit even smaller correlations or R-squareds. The source of data was the same as in the previous footnote.

TABLE VII
Association between Bank Structure and Local Economic Growth, Controlling for Banking Entry and State's Growth Rates

Variable	Model 1	2	3	4	5	6
Intercept	−18.29	−18.97	−24.57	−24.23	−26.99	−25.03
	(−7.15)*	(−7.06)*	(−10.12)*	(−10.90)*	(−6.92)*	(−7.16)*
log(#bks79)	3.25	—	1.86	—	2.00	
	(3.97)*		(2.52)**		(2.13)**	
log(#bks84)	—	3.50	—	1.94	—	1.86
		(3.78)*		(2.74)*		(2.07)**
Population	0.000505	0.000414	—	—	5.61×10^{-4}	6.29×10^{-4}
	(0.67)	(0.49)			(0.69)	(0.79)
% College	0.634	0.629	0.631	0.578	0.628	0.631
	(5.82)*	(5.72)*	(5.98)*	(6.45)*	(5.85)*	(5.88)*
Total deposits (1984, $mill.)	—	—	5.69×10^{-5}	5.62×10^{-5}	—	—
			(1.73)***	(1.73)***		
Net entry by banks (1979–1984)	−0.161	−0.249	—	—	1.59	−0.123
	(−1.71)***	(−2.42)**			(0.79)	(−0.07)
State growth rate[a] (1980–1990)	—	—	0.653	0.710	0.670	0.673
			(7.93)*	(9.73)*	(8.12)*	(8.17)*
# Observations	237	237	205	251	205	205
Adj. R-squared	0.256	0.257	0.430	0.451	0.427	0.426

Note. Dependent variable: Percentage growth in median money income per household by MSA, 1979–1989. (Heteroskedastic-consistent t-statistics in parentheses.)

Sources. 1996 County and City Extra: Annual Metro, City and County Data Book (Bernan Press, Lanham, Maryland); FDIC *Summary of Deposits; Statistical Abstract of the United States, 1995* (U.S. Bureau of the Census, Washington, D.C.).

* Significant at the 0.01 level.

** Significant at the 0.05 level.

*** Significant at the 0.10 level.

[a] State growth rate = percentage increase in personal income per capita (constant dollars), 1980–1990 in the MSA's state.

The negative coefficient on net entry invites further explanation. The high chargeoff rate of *de novo* banks (Tables I and II) may tend to impair economic growth. Then, even if banks exit from declining markets, the net effect of structural change across the sample could take the observed sign if there are more *de novo* entrants than exiting banks. A direct test of this effect would require data on gross, rather than net, entry and exit. Lacking that, the effect can be partially tested by estimating separate coefficients on positive versus negative values of entry.[26] Such regressions (not reported

[26] Nationwide, the structure of banking was fairly stable from 1979 through 1984, with a net increase from 14,364 to 14,496 insured commercial banks. Over this period there were 1676 new charters, 1462 unassisted mergers, 83 charter conversions from thrifts to banks, 182 failures, and 12 other exiting banks (source: FDIC's *Historical Statistics on Banking, 1934–1994*).

in the tables) indicated that positive net entry was not significantly associated with income growth, while net exit was significantly and *positively* associated with income growth (i.e., MSAs with sharper declines in the number of banks had higher income growth on average). This last result might be due to the fact that much "exit" during this period consisted of mergers, some of which were prompted by growing rather than shrinking market opportunities.

Such an explanation must recognize that both banking structure and local economic growth are affected by the state's macroeconomy and regulatory environment. Jayaratne and Strahan (1996) have found evidence that state laws restricting bank branching do affect economic growth rates as well as bank structure. Accordingly, Table VII also reports regressions controlling for each state's percentage growth rate of personal income per capita (constant dollars) from 1980–1990, as a proxy for the net effect of state-specific business conditions and regulations. MSAs spanning more than one state are omitted from these regressions. If the number of banks is merely a proxy for state-specific economic and regulatory effects, then including state growth rates should render the number of banks insignificant. As shown in Table VII, adding state growth rates nearly doubles the adjusted R-squared and somewhat reduces the magnitude of the coefficient on the log number of banks. However, the latter variable remains significant at the 0.05 level or better in each of these last four regressions. The bank entry variable, on the other hand, loses its significance in the presence of state growth rates. It should be noted that individual MSA growth rates are themselves components of the statewide growth rate, inducing some positive correlation between the two variables unrelated to exogenous causal factors.[27]

Other regressions not reported in the tables included the number of banks per capita in each MSA, a measure of banking density. Recent theories of the impact of financial intermediation on economic growth suggest that this measure might dominate the raw number of banks as a determinant of growth, whereas adverse borrower selection predicts a linkage between aggregate lending and the raw number of banks. Banks per capita was marginally significant—but with a *negative* coefficient—controlling for bank structure as of 1979; it was not significant controlling for bank structure as of 1984. The inclusion of per capita banks did not

[27] Restricting MSAs to those with growth rates between −20% and 20% removed around 20 to 30 observations from the full samples. The resulting estimates had roughly half the adjusted R-squareds of the full samples and smaller coefficients on the intercept, log number of banks, and education variables. In addition, total deposits and banks' entry rates ceased to be significant. Nevertheless, the coefficients on the log number of banks exceeded 2 and remained significant at the 0.01 level or better.

materially alter the magnitude or significance of the estimated coefficients on the other variables.[28]

In general, it appears to be a very 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 in Table V, the stronger association seems to be with subsequent growth rather than contemporaneous growth: the bank numbers for 1979 generated larger adjusted R-squareds, generally larger point estimates, and generally larger t-statistics than the bank numbers for 1984 in corresponding models (comparing models 1 and 5 and models 3 and 7 in Table V). In the semilog models, the regressions incorporating 1979 bank structure data exhibited slightly larger adjusted R-squareds, although 1984 bank structure figures yielded larger point estimates and t-ratios (comparing models 2 and 6 and models 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 III. One alternative explanation might be the stimulus provided by a net deposit insurance subsidy to the banking industry and thus indirectly to the communities served. However, the existence of such a subsidy has been sharply debated (Whalen, 1997), and in any case its effect should be reflected as much by total deposits as by the number of banks.

Another alternative explanation could be the stimulus to overall investment and growth if lending occurs at more competitive interest rates in less concentrated markets. But monopsony power on the deposit side, together with adverse borrower selection, could tend to offset this effect. Hannan and Liang (1993) and others have found evidence that deposit interest rates are higher in less concentrated markets. The higher loan loss rate theoretically and empirically associated with unconcentrated markets mandates a higher breakeven interest rate spread between loans and deposits in those markets. Thus, only if the level of excess profits declines sharply with increasing numbers of banks could market power dominate the sign of this linkage.

Stronger competition might also engender 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. Regardless of any additional

[28] As further tests of robustness, the models were estimated with additional control variables (not shown in the tables) including public expenditures per capita, the percentage of Democratic votes in the 1992 presidential election (two variables suggested by Jayaratne and Strahan, 1996), and the end-of-period unemployment rate in each MSA (a measure of resource utilization). These variables were not statistically significant and did not materially alter the magnitude or significance of the estimated coefficients on the other variables.

or alternative factors at work, Tables III and V together suggest 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 article has explored linkages among bank structure, loan performance, and local economic growth. Building on theoretical models of Broecker (1990) and Nakamura (1993), I have developed a theoretical framework that identifies additional dimensions of an adverse borrower selection effect and have also provided the first empirical tests of that effect. The linkages between chargeoff rates and market structure predicted by the theory are supported by the data. The empirical 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. This result parallels previous findings at the state level (Jayaratne and Strahan, 1996; Krol and Svorny, 1996) and across countries (King and Levine, 1993; Levine, 1998; Rajan and Zingales, 1998). Available data did not permit a precise explanation of this effect, but I 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 aggregate bank profits is consistent with this interpretation. A negative association between banks' entry and communities' income growth appears to rule out reverse causality; nevertheless, even if reverse causality is present, the robust linkage between the number of banks and the rate of income growth suggests that any adverse effects of larger numbers of banks (including higher loan chargeoff rates) do not in practice offset the benefits.

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 (1998). 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.

The findings here raise a number of issues for further exploration. The scope of corrective mechanisms can be examined more comprehensively; only common filters were analyzed here as a theoretical solution to adverse selection, while the empirical results for mature banks loosely suggest that certain other factors such as collateral and relationship lending may be effective as well. Signals were modeled here as costless, whereas in practice banks have a choice of screening technologies at various costs; endogenizing the tradeoff between signal precision and cost may yield further interesting findings. Our model suggests that competition among banks for the right of first refusal should be important; Broecker (1990) has analyzed this question for price competition, but nonprice competition such as marketing and pre-approved loan solicitations has not received formal analysis. It is possible that such analysis might support a long-standing concern among bankers and regulators that the banking industry may be peculiarly susceptible to "cut-throat competition" requiring corrective regulatory measures such as examinations for safety and soundness, restrictions on entry, and (as in former years) pricing restrictions.

REFERENCES

Altman, E. I., and Haldeman, R. (1995). Corporate credit-scoring models: Approaches and tests for successful implementation, *J. Commercial Lending* **78,** 10–22.

Asch, L. (1995). How the RMA/Fair, Isaac credit-scoring model was built, *J. Commercial Lending* **78,** 10–16.

Avery, R. B., Bostic, R. W., Calem, P. S., and Canner, G. B. (1996). Credit risk, credit scoring, and the performance of home mortgages, *Fed. Reserve Bull.* **82,** 621–648.

Avery, R. B., and Gordy, M. (1995). Interim report on the loan seasoning project, Federal Reserve Board of Governors, mimeo.

Berger, A. N. (1991). "The Profit–Concentration Relationship in Banking," Finance and Economics Discussion Series No. 176, Federal Reserve Board of Governors.

Berger, A. N. (1989). The price–concentration relationship in banking, *Rev. Econ. Statist.* **71,** 291–299.

Berger, A. N., and Humphrey, D. B. (1992). Megamergers in banking and the use of cost efficiency as an antitrust defense, *Antitrust Bull.* **37,** 541–600.

Berkovec, J. A., Canner, G. B., Gabriel, S. A., and Hannan, T. H. (1996). Mortgage discrimination and FHA loan performance, *Cityscape: A Journal of Policy Development and Research* **2,** 9–24.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational processes, *J. Polit. Econ.* **100,** 992–1026.

Bresnahan, T. F. (1982). The oligopoly solution concept is identified, *Econ. Letters* **10,** 87–92.

Broecker, T. (1990). Credit-worthiness tests and interbank competition, *Econometrica* **58,** 429–452.

DeYoung, R., and Hasan, I. (1998). The performance of de novo commercial banks: A profit efficiency approach, *J. Banking Finance* **22,** 565–587.

Easterly, W., Kremer, M., Pritchett, L., and Summers, L. H. (1993). Good policy or good luck? Country growth performance and temporary shocks, *J. Monet. Econ.* **32,** 459–483.

Emmons, W. (1993). Increased risk-taking versus local economic conditions as causes of bank failures, *Proc. Conference on Bank Structure and Competition*, Federal Reserve Bank of Chicago, 189–209.

Ferguson, M. F., and Peters, S. R. (1995). What constitutes evidence of discrimination in lending? *J. Finance* **50,** 739–748.

Galster, G. (1996). Future directions in mortgage discrimination research and enforcement, *in* "Mortgage Lending, Racial Discrimination, and Federal Policy" (J. Goering and R. Wienk, Eds.), pp. 679–716. Urban Institute Press, Washington, DC.

Hannan, T. H., and Liang, N. J. (1993). Inferring market power from time-series data: The case of the banking firm, *Int. J. Ind. Organ.* **11,** 205–218.

Hughes, J. P., Lang, W., Mester, L. J., and Moon, C.-G. (1996). Efficient banking under interstate branching, *J. Money, Credit, Banking* **28,** 1045–1071.

Interagency Task Force on Fair Lending (1994). "Policy Statement on Discrimination in Lending," Washington, D.C.

Jayaratne, J., and Strahan, P. E. (1996). The finance–growth nexus: Evidence from bank branch deregulation, *Quart. J. Econ.* **111,** 639–670.

King, R., and Levine, R. (1993). Finance and growth: Schumpeter might be right, *Quart. J. Econ.* **108,** 717–738.

Krol, R., and Svorny, S. (1996). The effect of the bank regulatory environment on state economic activity, *Reg. Sci. Urban Econ.* **26,** 531–541.

Ladd, H. F. (1998). Evidence on discrimination in mortgage lending, *J. Econ. Perspect.* **12,** 41–62.

Lau, L. J. (1982). On identifying the degree of competitiveness from industry price and output data, *Econ. Letters* **10,** 93–99.

Levine, R. (1998). The legal environment, banks, and long-run economic growth, *J. Money, Credit, Banking* **30,** 596–613.

Maddala, G. S. (1987). Limited dependent variable models using panel data, *J. Human Res.* **22,** 311–338.

Mester, L. J. (1993). Efficiency in the savings and loan industry, *J. Banking Finance* **17,** 267–286.

Nakamura, L. I. (1993). "Loan Screening Within and Outside of Customer Relationships," Federal Reserve Bank of Philadelphia Working Paper No. 93–15, June.

Nakamura, L. I., and Shaffer, S. (1993). Optimal screening by journals, Federal Reserve Bank of Philadelphia, mimeo.

Nathan, A., and Neave, E. H. (1989). Competition and contestability in Canada's financial system: Empirical results, *Can. J. Econ.* **22,** 576–594.

Neave, E. H., and Nathan, A. (1991). Reply to Perrakis, *Can. J. Econ.* **24,** 733–735.

Rajan, R. G., and Zingales, L. (1998). Financial dependence and growth, *Amer. Econ. Rev.* **88,** 559–586.

Shaffer, S. (1985). Competition, economies of scale, and diversity of firm sizes, *Appl. Econ.* **17,** 467–476.

Shaffer, S. (1989). Competition in the U.S. banking industry, *Econ. Lett.* **29,** 321–323.

Shaffer, S. (1993a). A test of competition in Canadian banking, *J. Money Credit Banking* **25,** 49–61.

Shaffer, S. (1993b). Can megamergers improve bank efficiency? *J. Banking Finance* **17,** 423–436.

Shaffer, S. (1996a). Evidence of discrimination in lending: An extension, *J. Finance* **51,** 1551–1554.

Shaffer, S. (1996b). Viability of traditional banking activities: Evidence from shifts in conduct and excess capacity, *Int. J. Econ. Business* **3,** 125–143.

Shaffer, S. (1999). "The Competitive Impact of Disclosure Requirements in the Credit Card Industry," *J. Reg. Econ.* **15,** 183–198.

Shaffer, S., and DiSalvo, J. (1994). Conduct in a banking duopoly, *J. Banking Finance* **18,** 1063–1082.

Stiglitz, J. E., and Weiss, A. (1981). Credit rationing in markets with imperfect competition, *Amer. Econ. Rev.* **71,** 393–410.

Thakor, A. V. (1996). Capital requirements, monetary policy, and aggregate bank lending: Theory and empirical evidence, *J. Finance* **51,** 279–324.

Whalen, G. (1997). "The Competitive Implications of Safety Net-Related Subsidies," Working Paper 97–9, Office of the Comptroller of the Currency, May.