

B

# ECONOMIC REVIEW

1998 Quarter 4

---

**Beneath the Rhetoric: 2**  
**Clarifying the Debate on**  
**Mortgage Lending Discrimination**

by Stanley D. Longhofer and Stephen R. Peters

---

**Banking and Commerce: 14**  
**How Does the United States**  
**Compare to Other Countries?**

by João A.C. Santos



**FEDERAL RESERVE BANK  
OF CLEVELAND**

## 1998 Quarter 4

---

**Beneath the Rhetoric:  
Clarifying the Debate on  
Mortgage Lending Discrimination** 2  
by Stanley D. Longhofer and Stephen R. Peters

---

**Banking and Commerce:  
How Does the United States  
Compare to Other Countries?** 14  
by João A.C. Santos

## **Beneath the Rhetoric: Clarifying the Debate on Mortgage Lending Discrimination**

by Stanley D. Longhofer and Stephen R. Peters

2

The authors' simple model of the mortgage underwriting process provides a framework within which to define discrimination and various notions of the default rate. By providing those with differing views a common framework for discussing their positions, the model clarifies and reconciles some of the most controversial issues in the debate over mortgage discrimination. It also shows how this theoretical framework can help in the design of practical policy responses to this vexing social problem.

## **Banking and Commerce: How Does the United States Compare to Other Countries?**

by João A.C. Santos

14

This article begins with a discussion of the affiliation between banking and commerce throughout American history. It finds that there were important occasions when banks were allowed to invest in nonfinancial firms and that restrictions on firms' investments in banks are a recent phenomenon. In comparing current U.S. regulations on the affiliation between banks and firms to those in force abroad, the article shows that foreign countries have much more liberal banking laws. The analysis of banks' investments in equities in those countries, however, shows that such investments represent a small fraction of banks' assets. The article ends with a brief review of the literature on the implications of banks' investments in firms.

*Economic Review* is published quarterly by the Research Department of the Federal Reserve Bank of Cleveland. Copies of the *Review* are available through our Corporate Communications & Community Affairs Department. Call 1-800-543-3489 (OH, PA, WV) or 216-579-2001, then immediately key in 1-5-3 on your touch-tone phone to reach the publication request option. If you prefer to fax your order, the number is 216-579-2477.

*Economic Review* is available electronically through the Cleveland Fed's site on the World Wide Web: <http://www.clev.frb.org>. We also offer a free online subscription service to notify readers of additions to our Web site. To subscribe, please send an e-mail message to [econpubs-on@clev.frb.org](mailto:econpubs-on@clev.frb.org).

Editors: Michele Lachman  
Lisa McKenna  
Deborah Zorska  
Design: Michael Galka  
Typography: Liz Hanna

Opinions stated in *Economic Review* are those of the authors and not necessarily those of the Federal Reserve Bank of Cleveland or of the Board of Governors of the Federal Reserve System.

Material may be reprinted provided that the source is credited. Please send copies of reprinted material to the editor.

# Beneath the Rhetoric: Clarifying the Debate on Mortgage Lending Discrimination

by Stanley D. Longhofer and Stephen R. Peters

Stanley D. Longhofer is an economist at the Federal Reserve Bank of Cleveland and Stephen R. Peters is a visiting assistant professor of finance at the University of Illinois at Urbana–Champaign.

## Introduction

Home ownership has long been identified as a central element of the “American Dream”; in fact, many treat the two as synonymous. No wonder, then, that much of federal anti-discrimination law focuses on housing and the lending markets used to finance home purchases. For more than 30 years, bank regulators have been charged with the task of rooting out illegal discrimination in credit markets.

The long-standing debate over whether lenders systematically discriminate against minorities received added momentum in 1992 with the release of the now-famous “Boston Fed Study” (Munnell et al. [1992]), which purported to find that minority applicants in the Boston area are roughly half again as likely as similarly situated white applicants to be denied mortgage loans. Although their conclusions have been hotly debated in the ensuing years, there can be no question that Munnell et al. put the mortgage discrimination debate on the front burner in both policy and academic circles.

This debate still rages today. Yet after more than six years, we are not much closer to knowing whether discrimination is a widespread problem in mortgage markets. Indeed,

economists have yet to agree on whether it is more effective to detect discrimination using denial rate analyses like those employed by Munnell et al. or by examining borrower default rates. Nor have we settled on the appropriate policy response to make if discrimination is in fact as prevalent as Munnell et al. suggest.

A key roadblock in answering these questions has been the lack of a solid theoretical foundation on which to conduct the debate. Without a consistent economic theory of the mortgage application, underwriting, and default processes, fundamental tasks such as defining discrimination and outlining how it might be detected have proven slippery. As a consequence, more advanced issues such as identifying the source of discrimination and evaluating potential policy responses have been obviated, despite longstanding calls to “move beyond a debate over whether discrimination exists to a discussion of how best to eradicate it” (Galster, [1993, p. 146]).

In this article, we outline a simple theoretical model of the mortgage underwriting process, originally developed by Longhofer and Peters (1999), which provides a framework for clearly defining discrimination and various notions of the default rate. By giving those with differing

views a common framework for arguing their positions, this model clarifies and reconciles some of the most contentious questions at the heart of the controversy over mortgage discrimination. We also show how this theoretical foundation can aid in designing practical policy responses to this vexing social problem.

The next section reviews the highlights of the debate that ensued in the wake of the initial study by Munnell et al. Section II introduces the Longhofer–Peters model of mortgage underwriting and shows how it can be used to clarify many of the definitions and assumptions that underlie this debate. Section III reexamines the mortgage discrimination debate in light of that model, and section IV concludes.

## I. The Mortgage Discrimination Debate

The debate over whether lenders systemically discriminate against minority applicants began in earnest with the release of Munnell et al. (1992).<sup>1</sup> The authors of this study found that minority applicants in the Boston metropolitan area were approximately 50 percent more likely to be rejected than whites, even after controlling for the factors that banks use to determine an applicant's creditworthiness. "This means that 17 percent of black or Hispanic applicants instead of 11 percent would be denied loans, even if they had the same obligation ratios, credit history, loan to value, and property characteristics as white applicants. In short, the results indicate that a serious problem exists in the market for mortgage loans, and lenders, community groups, and regulators must work together to ensure that minorities are treated fairly" (Munnell et al. [1992, p. 44]).

In his Nobel Lecture, Becker (1993b) challenges these findings. Recasting his economic theory of discrimination in the context of the mortgage lending market, Becker argues that bigoted lenders are willing to sacrifice the profit they might earn by approving marginally qualified minority applicants in order to satisfy their "tastes for discrimination." Thus, the proper way to detect discrimination is to do so directly, by comparing the relative profitability of loans to minorities and whites: "This requires examining the default and other payback experiences of loans, the interest rates charged, and so forth. If banks discriminate against minority applicants, they should earn *greater* profits on the loans actually made to them than on those to whites. The reason is that discriminating banks would

be willing to accept marginally profitable white applicants" (Becker [1993b, p. 389]).

If banks discriminate against minority borrowers, Becker argues, we should observe minorities defaulting less frequently than whites on average. "[T]he theory of discrimination contains the paradox that the rate of default on loans approved for blacks and Hispanics by discriminatory banks should be lower, not higher, than those on mortgage loans to whites. The reason is that such banks only accept the very best minority candidates" (Becker [1993a]). Because discriminators who are following their tastes will pass up better-qualified minority applicants for less-qualified white ones, the average creditworthiness of the minorities who actually receive loans will be higher than that of whites.

Macey (1994) uses Becker's argument to refute the principal conclusions of Munnell et al., citing that study's own data.<sup>2</sup> "The default rates for white and black mortgage loan applications are equal across census tracts. If bankers were discriminating by turning down marginally qualified black applicants while accepting marginally qualified white applicants, then default rates among whites would be higher. But bankers are accepting the same level of risk from both black and white applicants" (Macey [1994]).

The validity of this argument regarding average default rates relies on the assumption that white and minority credit risk is equally distributed in the borrower pool. In fact, however, Munnell et al.'s data suggest that minority applicants (and borrowers) are less creditworthy than whites on average (see Munnell et al. [1996], table 1).

Galster (1993) uses this fact to argue that minorities will default more often than whites in the absence of discrimination. Thus, he concludes that equal average default rates across minority and white borrower pools imply taste-based discrimination against minorities. On these grounds, Galster suggests that default rates cannot be considered a reliable indicator of discrimination.<sup>3</sup>

The above arguments revolve around whether average default rates can be used to uncover discriminatory behavior on the part of lenders. Others, including Calomiris et al. (1994)

■ 1 Later revised and published as Munnell et al. (1996).

■ 2 See also Brimelow and Spencer (1993) and Becker (1993a).

■ 3 Tootell (1993), Browne and Tootell (1995), and Munnell et al. (1996) echo Galster's conclusions.

and Ferguson and Peters (1995) focus on the interpretation of *marginal* default rates. These authors argue that, regardless of the relative distributions of creditworthiness across racial groups, in the absence of taste-based discrimination the marginal borrower—the least creditworthy applicant to be approved for a loan—will have the same creditworthiness across groups. That is, if banks hold all applicants to the same credit standard, the marginal default rate will be the same for all racial groups. If this is true, then comparisons of the marginal default rate across races may enable regulators to detect taste-based discrimination (in the manner of Becker [1993b]).

In summary, a variety of apparently conflicting tests have been proposed for inferring whether lenders discriminate against minority applicants: Lower average default rates for minorities are evidence of discrimination against minorities (for example, Becker [1993b]); equal average default rates for minorities and whites are evidence of discrimination against minorities (Galster [1993]); and, lower marginal minority default rates are evidence of discrimination against minorities (Calomiris et al. [1994]). Which of these conclusions are we to believe? How should we go about uncovering illegal discrimination? In the next section, we develop a simple theoretical model that can help answer these questions.

## II. A Theoretical Underpinning for the Debate

In order to define what constitutes discrimination and determine how such behavior might be detected, we must first understand how mortgage applications are underwritten. To do this, we use the mortgage underwriting model developed in Longhofer and Peters (1999).

Consider a world in which individuals want to buy a house, but lack sufficient funds to do so. As a result, they must obtain loans from a financial institution, which we will call a “bank.” Individuals in this world are distinguished by two characteristics. The first is the likelihood with which they repay their loans, which we denote by  $\theta \in [0, 1]$ . The second is their membership in one of two “groups,”  $A$  or  $B$ . For ease of exposition, we will often refer to group  $A$  as the “white” group and to group  $B$  as the “minority” group. It should be clear, however, that these groups may alternatively be thought of as distinguishing individuals according to any publicly observable characteristic over which fair-lending laws might prohibit discrimination.

Let  $g_A(\theta)$  represent the density of creditworthiness in the group  $A$  applicant pool and  $g_B(\theta)$  represent the corresponding density for members of group  $B$ . It is important to note that these two densities generally will differ from one another.<sup>4</sup> We assume that both  $g_A(\theta)$  and  $g_B(\theta)$  are known to all banks.

Based on their costs of funds and the competitive interest rate in the market, banks determine a  $\theta^*$  defining the minimum acceptable creditworthiness they are willing to approve. In a full-information world, lenders would know each individual’s true  $\theta$  and would approve only those applicants with  $\theta \geq \theta^*$ . Unfortunately, lenders cannot perfectly observe  $\theta$ , but instead observe a signal  $s$  that is correlated with  $\theta$ . This signal can be thought of as a summary of all the information a lender collects on a loan application, including the applicant’s current income ratios and past credit history, the characteristics of the subject property, and so on.<sup>5</sup>

Let  $p(s|\theta)$  denote the likelihood that a lender observes signal  $s$  from an applicant of type  $\theta$ .<sup>6</sup> Using this signal generation process and the distribution of creditworthiness in the applicant pool, lenders derive an estimate of an applicant’s expected creditworthiness, given his signal. We will often refer to this expected creditworthiness as an applicant’s inferred “quality,” which is denoted by

$$(1) \quad q_i(s) = \int_{T_i} \frac{p(s|\theta)g_i(\theta)}{\omega_i(s)} d\theta, \quad i = A, B$$

where  $\omega_i(s) = \int_{T_i} p(s|\theta)g_i(\theta) d\theta$ ,  $i = A, B$ , is the density of signals received from members of group  $i$ , and  $T_i$  is the set of group  $i$  applicants that apply for loans:  $\{\theta | g_i(\theta) > 0\}$ .

Expression (1) makes it clear that an applicant’s inferred quality is simply his expected creditworthiness, where this expectation is taken over the lender’s (Bayesian) posterior beliefs about the applicant’s creditworthiness. Clearly,

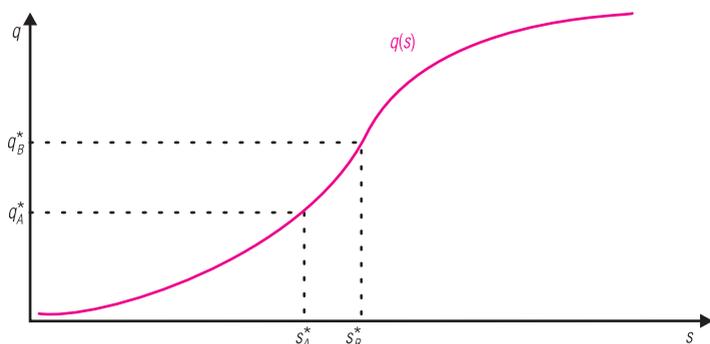
■ 4 Longhofer and Peters (1999) analyze how individual self-selection can lead to endogenous differences in credit risk across groups. Alternatively, one could imagine that the distributions of credit risk in the underlying population differs across groups, leading to differences in these densities for the group applicant pools even in the absence of self-selection (see, for example, Calomiris and Stutzer [1995]).

■ 5 Although an applicant’s  $\theta$  is assumed to be private information, the group to which he belongs can be costlessly observed by lenders, giving them an opportunity to discriminate on this basis if they so choose.

■ 6 Although we assume that this signal generation process is the same for members of both groups, one could imagine a world in which it differed. Such a difference in  $p(s|\theta)$  across groups lies at the heart of the cultural affinity hypothesis proposed by Calomiris et al. (1994).

FIGURE 1

### Taste-Based Discrimination (Bigotry)



Bigoted lenders act on a taste for discrimination by holding minority applicants to a higher creditworthiness threshold ( $q_A^* < q_B^*$ ). This implies that bigoted lenders will hold minorities to a more stringent underwriting standard than they do whites ( $s_A^* < s_B^*$ ).

this posterior will be determined by both the distribution of credit risk in the overall applicant pool and the specific signal sent by the applicant. Given these beliefs, a lender will approve an applicant if and only if  $q(s) \geq q^* \equiv \theta^*$ .

Longhofer and Peters (1999) show that an applicant's inferred creditworthiness is increasing in his signal. That is,  $q'(s) > 0$ , confirming the intuition that applicants who send better signals tend to be more creditworthy. More importantly, this result assures us that there exists a unique cutoff signal,  $s^*$ , such that all applicants sending signals better than  $s^*$  are approved while those sending worse signals are not.

### Defining Discrimination

The most striking omission in the debate over mortgage discrimination has been the lack of a formal definition of what behavior actually constitutes "discrimination." Perhaps this should not be surprising, given that the term is not defined in either the Equal Credit Opportunity Act or the Fair Housing Act—the two laws that directly prohibit discriminatory practices in the mortgage underwriting process. Presumably, this omission indicates Congress' belief that clarification would be superfluous because the meaning of the term is well understood by all reasonable people.

In our opinion, however, much of the controversy over how best to detect discrimination arises precisely because of differing implicit definitions of this central concept. To clarify the

law's intent, we turn to *Merriam Webster's Collegiate Dictionary*, 10th ed., which defines discrimination as "the act, practice, or an instance of discriminating [making a distinction] categorically rather than individually." In other words, any difference in treatment across individuals based solely on group membership—rather than on personal characteristics specifically related to the performance of the loan—would constitute discriminatory behavior under the law. It is our interpretation (and the practice of the major bank regulators) that such a definition precludes lenders from applying group-specific underwriting standards, even if group membership is correlated with loan performance.<sup>7</sup>

Consistent with this idea, consider the following:

**Definition:** A lender is said to **discriminate against minority applicants** if it requires them to meet a more stringent underwriting standard than it does white applicants; i.e., if  $s_B^* > s_A^*$ .

When discrimination is explicitly defined in this way, it becomes clear that lenders' incentives to discriminate can arise for two distinct reasons. First, lenders may have Beckerian "tastes for discrimination," bigoted preferences that would manifest themselves through differences in the  $q^*$  required from members of each group. In other words, we say that a lender exhibits bigoted preferences if  $q_B^* > q_A^*$ . This source of discrimination is depicted in figure 1, which assumes that  $q_A(s) = q_B(s)$  for every  $s$ .<sup>8</sup> In this case, the monotonicity of  $q$  implies that a bigoted lender will have an incentive to discriminate against members of group  $B$  by setting  $s_B^* > s_A^*$ .

Even without such tastes for discrimination, however, lenders may still have an incentive to discriminate if the overall pool of minority applicants is less creditworthy on average than the white applicant pool. In other words, if the minority applicant pool consists of relatively more low- $\theta$  individuals,  $q_A(s)$  will lie above  $q_B(s)$  for every  $s$ , giving banks an incentive for *statistical discrimination*.<sup>9</sup>

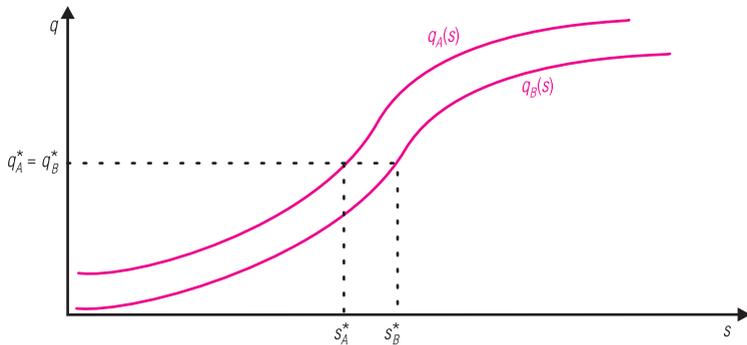
■ 7 A noteworthy exception is the Equal Credit Opportunity Act, which specifically allows age to be considered in credit-scoring models, but only if it is to the advantage of older applicants.

■ 8 This requires that the distribution of credit risk,  $g(\theta)$ , be the same across the two applicant pools.

■ 9 Strictly speaking, it is only necessary that  $q_A(s) > q_B(s)$  for  $s < s_B^*$ ; Longhofer and Peters (1999) present a case in which statistical discrimination arises and yet  $q_A(s) < q_B(s)$  for large  $s$ .

FIGURE 2

## Statistical Discrimination



If minority applicants are less creditworthy on average than whites, this will be reflected in the lender's assessment of an applicant's inferred quality. Thus,  $q_A(s)$  will lie above  $q_B(s)$ . In this case, even if lenders hold white and minority applicants to the same credit standard,  $q_A^* = q_B^*$ , they will have an incentive to statistically discriminate against minorities by setting  $s_A^* < s_B^*$ .

Statistical discrimination is depicted by figure 2, in which lenders set the same minimum creditworthiness standard for both white and minority applicants. Nevertheless, because minority applicants are less creditworthy (have lower  $\theta$ s) on average, lenders *believe* that a minority applicant sending any signal  $s$  is less creditworthy than a white applicant sending the same signal. Put another way, given the distribution of credit risk in their applicant pools, lenders know that a minority applicant sending any signal  $s$  is more likely to default than a white applicant sending the same signal. As a result, lenders have an incentive to hold minority applicants to a higher  $s^*$ .

Of course, these beliefs do not arise in a vacuum. As long as lenders accurately estimate the distribution of credit risk in their applicant pools,  $g_i(\theta)$ , the probability that a group  $i$  applicant repays his loan is  $q_i(s)$ . Thus, lenders that statistically discriminate find their beliefs validated *ex post*; on average, minority borrowers sending signal  $s_B^*$  repay their loans at exactly the same rate as white borrowers sending signal  $s_A^*$ , and both repay at the minimum rate acceptable to the bank.

## Defining Default Rates

Given the above discussion, we can now define and analyze the different notions of default rates that have arisen in the mortgage discrimination controversy. As it turns out, participants in the debate rarely use the most useful of these notions, the *conditional default rate*.

Definition: The **conditional default rate** for members of group  $i$  is the fraction of group  $i$  borrowers that actually default, conditional on their signal  $s$ :  $d_i(s) \equiv 1 - q_i(s)$ ,  $i = A, B$ .

As discussed above, the likelihood that a group  $i$  borrower with signal  $s$  will end up defaulting on his loan is  $1 - q_i(s)$ . If there is a sufficiently large population of borrowers, then the actual fraction of group  $i$  borrowers sending signal  $s$  that default will be  $d_i(s)$ . From a lender's perspective, this conditional default rate is what matters most. After all, the lender's underwriting decision is ultimately determined by the likelihood that an applicant will default, given the information on his application. This is exactly what  $d_i(s)$  measures.

Most of the mortgage discrimination debate, however, has focused on other notions of default rates.

Definition: The **average default rate** is the fraction of all borrowers who actually default on their loans:

$$(2) \quad \bar{d}_i \equiv \int_{s \geq s_i^*} d_i(s) \frac{\omega_i(s)}{\int \omega_i(s) ds} ds, \quad i = A, B.$$

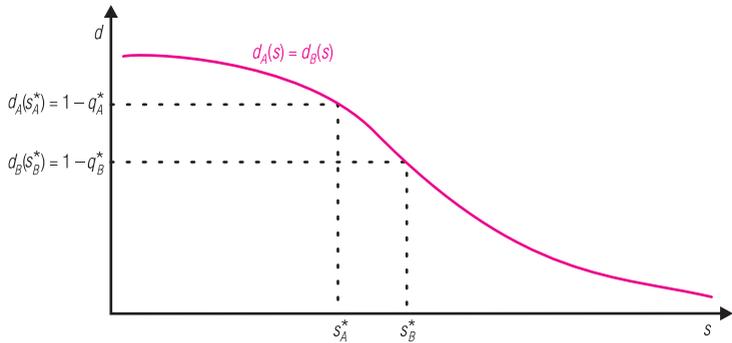
Thus, the average default rate is simply the weighted average of the conditional default rates of individuals who were approved, where the weights are determined by the distribution of signals sent by approved applicants.

Finally, Berkovec et al. (1994), Calomiris et al. (1994), and Ferguson and Peters (1997, 1998) all consider varying notions of the "marginal" rate of default (or denial). Unfortunately, this concept has often been somewhat ill-defined.<sup>10</sup> It can be most precisely defined as follows:

Definition: The **marginal default rate** is the fraction of defaulters among the borrowers sending the lowest approved signal:  $d_i^m \equiv d_i(s_i^*) = 1 - q_i(s_i^*)$ ,  $i = A, B$ .

FIGURE 3

## Default Rates under Bigotry



If credit risk is identically distributed across the minority and white applicant pools, bigoted lenders will require minority applicants to meet a more stringent underwriting standard (a higher  $s^*$ ). Because the minority applicants who are rejected are those with the highest default risk, the average default rate will be lower for minorities than it will be for whites.



In the context of our model, the notion of a marginal default rate may be easily understood. In the real world, however, it is virtually impossible for regulators to identify the unique cut-off signal  $s^*$  used by lenders.<sup>11</sup> As a practical matter, then, the marginal default rate is often taken as either the average or the conditional default rate among borrowers sending signals below some arbitrary threshold. For example, Berkovec et al. (1994) analyze conditional default rates among FHA borrowers, under the presumption that they are riskier than those obtaining conventional loans; the authors further attempt to focus on “marginal” borrowers by dividing their sample into risk quartiles, comparing conditional default rates across white and minority borrowers in each quartile.

### III. Deciphering the Debate

In section II, we introduced a simple theoretical model of the mortgage underwriting process and used it to help define discrimination and various notions of the default rate. In the present section, we use these definitions to reconsider the various arguments outlined in section I. In particular, we analyze the accuracy and the consistency of these arguments and consider why this debate has been so difficult to resolve. Finally, we discuss how our model’s formal structure can be used to develop a better understanding of the underlying causes of any discrimination that does exist in the mortgage

market. As we will argue, understanding the root causes of the problem is essential to designing effective policy responses.

### The Default Rate Controversy

We begin by imagining a world where credit risk is identically distributed in both the minority and white applicant pools; that is,  $g_A(\theta) = g_B(\theta)$ ,  $\forall \theta$ . In such a world, the only possible source of discrimination is bigotry, which manifests itself as a higher required standard of creditworthiness for minorities ( $q_B^* > q_A^*$ ). As shown in figure 1, lenders with such a taste for discrimination will respond by setting  $s_B^* > s_A^*$ . Nevertheless, the fact that  $g_A(\theta) = g_B(\theta)$ ,  $\forall \theta$  implies that the inferred creditworthiness of a white applicant and a minority applicant, each of whom sends the same signal  $s$ , will be identical; that is  $q_A(s) = q_B(s)$ ,  $\forall s$ . Thus, conditional default rates will be the same across the two groups. But if conditional default rates are the same across groups, the average default rates of the two groups must diverge whenever banks set  $s_B^* > s_A^*$ , that is,  $\bar{d}_A > \bar{d}_B$ . Put another way, if minority and white borrowers are equally creditworthy on average, then any difference in the average default rate across groups must arise from lenders’ tastes for discrimination.

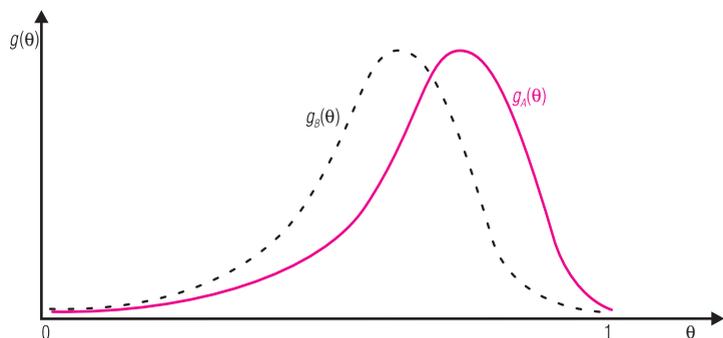
This idea is shown graphically in figure 3, in which conditional default rates are identical across the two groups. Because  $q_B^* > q_A^*$ , however, lenders set  $s_B^* > s_A^*$ , implying that white applicants with  $s \in [s_A^*, s_B^*]$  are approved, while minority applicants sending the same signals are not. Ex post, these borrowers are the least creditworthy and hence the most likely to default. Because these applicants are excluded from the minority borrower pool, the average default rate of minorities who are actually approved for loans must be lower than that of whites.

Pursuing this logic, Becker (1993a), Brimelow and Spencer (1993), and Macey (1994) claim that Munnell et al.’s own data refute a conclusion of discrimination. After all, they argue, default rates across census tracts do not appear to be systematically related to the percentage of minorities residing in those tracts. While not conclusive, this evidence seems inconsistent with the notion that minorities default more frequently on average than whites.

■ 11 For a discussion of how this model relates to econometric analyses of  $s^*$  by regulators and researchers, see Craig et al. (1998).

FIGURE 4

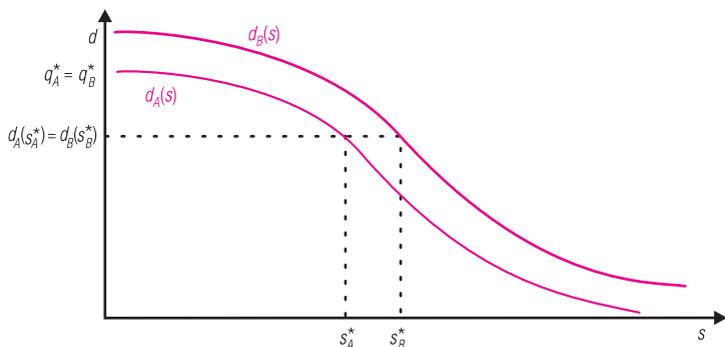
## Differences in Underlying Creditworthiness across Groups



If minority applicants are less creditworthy on average than white applicants, the density of credit risk in the minority applicant pool will lie to the left of the density for the white applicant pool.

FIGURE 5

## Statistical Discrimination and Default Rates



If minority applicants are less creditworthy than their white counterparts on average, lenders will have an incentive to statistically discriminate against them, even in the absence of bigotry. They do so by choosing  $s_A^*$  and  $s_B^*$  so as to equalize the inferred quality of the last applicant approved from each group [ $q_A(s_A^*) = q_B(s_B^*)$ ]. This in turn implies that the marginal default rate of the two groups will be identical. However, the relative average default rates will be ambiguous.

Given their assumptions, Becker (1993a), Brimelow and Spencer (1993), and Macey (1994) are all correct in their conclusions. In our model, however, it is clear that for a group of borrowers, the average default rate depends not only on the underwriting standard applied to that group, but also on the distribution of signals sent by those borrowers,  $\omega_i(s)$ . This, in turn, depends on the distribution of credit risk in the group's applicant pool. Thus, ex ante differences in credit risk across applicant pools can lead to ex post differences in observed

average default rates, even in the absence of discrimination.

This is the essential point made by Tootell (1993) and Galster (1993). They argue that Munnell et al.'s data verify that minority applicants are less creditworthy on average than their white counterparts.<sup>12</sup> This stylized fact is illustrated in figure 4, where  $g_B(\theta)$  is shifted slightly to the left of  $g_A(\theta)$ .<sup>13</sup> In the context of our model, such differences in credit risk across groups cause lenders to adjust their assessment of an applicant's inferred quality, so that  $q_A(s) > q_B(s)$ ,  $\forall s$ . As a result, if lenders were to hold both groups to the same  $s^*$ , the average default rate would be lower for white applicants. On the other hand, if lenders do set  $s_A^* < s_B^*$ , then the relative average default rate between the two groups will be ambiguous (see figure 5).

Thus we see that, if lenders are bigoted, there are two counteracting effects on minority default rates: the bigotry effect and lower average minority creditworthiness. On the one hand, lender tastes for discrimination imply that relatively more of the less-creditworthy white borrowers get included when calculating  $\bar{d}_A$ . On the other hand, for any given  $s$ , a minority borrower is more likely to default than a white borrower. Which of these two effects will dominate in practice is unclear.<sup>14</sup> As a result, Tootell and Galster both argue, average default rates cannot be used to disprove a charge of discrimination. In fact, if minority borrowers are less creditworthy on average than their white counterparts, a finding that both groups default at the same rate could only be consistent with lenders acting on bigotry against minorities.

Even though average default rates cannot disprove the existence of discrimination, other notions of the default rate may still be useful. The essential problem with average default rates is that they incorporate too much information. In contrast, conditional and marginal default rates focus on a specific subset of borrowers, making it easier to interpret what lies behind any differences across groups.

At its core, Becker's argument is driven by the fact that, if lenders hold minorities to a more stringent underwriting standard arising from bigotry, the "marginal" minority borrower

■ 12 The Federal Reserve's National Surveys of Consumer Finances also support this conclusion.

■ 13 Formally, one can think of the cumulative distribution function of credit risk in the white applicant pool,  $G_A(\theta)$ , as being first-order stochastic dominant over  $G_B(\theta)$ , the cumulative distribution of credit risk in the minority applicant pool; see Ferguson and Peters (1995).

■ 14 In addition, the distribution of signals,  $\omega(s)$ , will differ across groups, further complicating this calculation.

TABLE 1

### Relative Default Rates Implied by Bigotry and Statistical Discrimination

	Bigotry Only	Statistical Only	Both
Conditional default rates	$d_A(s) = d_B(s), \forall s$	$d_A(s) < d_B(s), \forall s$	$d_A(s) < d_B(s), \forall s$
Marginal default rates	$d_A(s_A^*) > d_B(s_B^*)$	$d_A(s_A^*) = d_B(s_B^*)$	$d_A(s_A^*) > d_B(s_B^*)$
Average default rates	$\bar{d}_A > \bar{d}_B$	$\bar{d}_A \geq \bar{d}_B$	$\bar{d}_A \geq \bar{d}_B$

(that is, the one with the lowest approved signal) will be more creditworthy than the marginal white borrower (see figure 3). In contrast, if lenders practice statistical discrimination, they do so precisely because they wish to ensure that the marginal white and marginal minority borrowers are equally creditworthy. In other words, statistical discrimination arises when lenders try to offset differences in the overall applicant pools across races (see figure 5). Thus, when we focus on marginal default rates, we see that Becker's argument remains valid, even in the presence of differential creditworthiness across racial groups.<sup>15</sup>

Unfortunately, regulators and econometricians cannot precisely observe a borrower's actual signal, nor can they pin down a lender's true underwriting guidelines. Thus, identifying the marginal borrower from each group is a practical impossibility. One way around this problem is to examine the default risk within a subset of borrowers whose creditworthiness is below some arbitrary threshold, as in Berkovec et al. (1994). Using a sample of FHA borrowers, they find that minorities in the highest-risk quartile have a significantly higher conditional default rate than do whites in the same risk class.<sup>16</sup> If we generically treat all borrowers in this risk class as "marginal," the results of Berkovec et al. would appear to suggest that statistical discrimination is the best explanation for those of Munnell et al.

Table 1 summarizes the default-rate implications discussed in this section. In a world in which members of both groups are equally creditworthy on average, conditional default rates will likewise be the same across groups [ $d_A(s) = d_B(s), \forall s$ ]. In this case, if lenders exhibit tastes for discrimination the default rate of the marginal minority borrower will be lower than

that of the marginal white borrower [ $d_A(s_A^*) > d_B(s_B^*)$ ], causing the average default rate of minority borrowers to be lower than that of whites ( $\bar{d}_A > \bar{d}_B$ ).

In contrast, when lenders have an incentive to statistically discriminate, minorities' conditional default rate will be higher than that of whites [ $d_A(s) < d_B(s), \forall s$ ]; average default rates will be ambiguous. Nevertheless, lenders acting solely on a profit motive will discriminate against minorities only to the extent necessary to equalize the marginal default rate of members of each group [ $d_A(s_A^*) = d_B(s_B^*)$ ].

Finally, when lenders act on both statistical and taste-based motives, conditional default rates will continue to be higher for minorities [ $d_A(s) < d_B(s), \forall s$ ]. Interestingly, however, lender bigotry will cause the default rate of the marginal minority borrower to fall below that of the marginal white borrower [ $d_A(s_A^*) > d_B(s_B^*)$ ], despite the higher conditional default rates for minority borrowers.

It is important to note that while conditional default rates may be useful in determining whether lenders have an incentive to discriminate, they cannot be used to determine whether lenders actually act on this incentive. Conditional default rates across groups will diverge whenever minority borrowers on average are substantially less creditworthy than white borrowers. This will be true whether or not lenders actually require minorities to meet a higher underwriting threshold,  $s^*$ . Thus, we must rely on denial rate analyses such as Munnell et al. to uncover discrimination and use default rate information to help determine the underlying source, if discrimination is determined to exist; we discuss the importance of this issue in more detail next.

### What Constitutes Discrimination?

Taken together, Munnell et al.'s denial rate findings and Berkovec et al.'s default rate results pose a puzzle: How can marginal minority borrowers be significantly more likely to default than their white counterparts, yet still be less likely to be approved in the first place?

■ 15 This point was made by Calomiris et al. (1994).

■ 16 Although Berkovec et al. do control for numerous property and personal characteristics, an important limitation of their analysis is that their data do not contain any information on individual credit histories. The authors attempt to correct for any bias resulting from this omission, and still conclude that marginal minority borrowers default at a significantly higher rate than do whites.

The answer is statistical discrimination. In contrast to the taste-based discrimination considered by Becker, statistical discrimination is based on lenders' beliefs about applicant creditworthiness. Calomiris et al. (1994), Calem and Stutzer (1995) and Longhofer and Peters (1999) each show that statistical discrimination can occur if lenders' beliefs about creditworthiness differ across groups.<sup>17</sup> For example, Longhofer and Peters (1999) show that if a lender believes that minority applicants' average creditworthiness is lower than that of white applicants, then minority applicants will have to clear a higher hurdle—that is, minorities will be required to have better applications than whites—in order to receive a loan. Thus, although the inferred creditworthiness (the posterior assessment of creditworthiness) of white and minority borrowers is the same at the margin, econometric tests that use a borrower's application information (the signal) to proxy for his true creditworthiness will find that minorities default more often.

Although statistical discrimination provides a straightforward, plausible resolution of the results of Munnell et al. and Berkovec et al., many economists appear to be uncomfortable with this explanation. In fact, they seem to operate under the assumption that statistical discrimination is not, or should not be, illegal.

Where would such an assumption come from? Current law permits lenders to include most any underwriting variable that is demonstrably related to the profitability of a loan.<sup>18</sup> For example, lenders may consider an applicant's past credit history, even though this variable is highly correlated with race, because it is also strongly associated with an applicant's likelihood of repaying the loan. On the basis of this observation, it is reasonable to conclude that the proper scope of fair-lending law is to prohibit discrimination that arises for motives that are not profit-based, that is, from bigotry.

For example, Becker and Macey's focus on average default rates would seem rather misplaced, considering that average minority creditworthiness is lower than average white creditworthiness. Such a focus is perfectly reasonable, however, if one believes that statistical discrimination is not, in fact, discrimination at all. We are also struck by the lengths to which Munnell et al. go to reject the notion that their findings are driven by statistical

discrimination: "The dearth of any evidence that minorities default more frequently, given their economic fundamentals, makes a conclusion of economically rational [statistical] discrimination problematic" (Munnell et al., [1996, p. 45]).

As we've already argued, however, existing fair-lending laws appear to prohibit both taste-based and statistical discrimination implicitly. Indeed, if statistical discrimination were permissible, lenders could simply include race as a variable in a statistically verified credit-scoring model; as long as the race variable showed a statistically significant impact on the loan's overall profitability, it would be permissible to consider race under such an interpretation of the law. Of course, this is expressly illegal. Furthermore, the "disparate impact test" in fair-lending enforcement effectively precludes the use of underwriting variables that themselves have no information content, but rather serve as proxies for race.

This discussion highlights why it is so important to confront issues like mortgage discrimination with a sound theoretical model. In the context of our formal model of the mortgage underwriting process, the definition of discrimination and the proper way of measuring it become clear. When both statistical discrimination and bigotry can result in differences in  $s^*$  across groups, the question of whether default rates can be used to detect discrimination becomes largely irrelevant.

This is not to suggest, however, that default rates are useless in the mortgage discrimination debate. In fact, one of the most important questions is how best to respond to any discrimination that does exist in the mortgage market. It is to this question that we turn next.

## Uncovering the Source of Discrimination

Although borrower default rates are poorly suited for determining whether or not lenders discriminate against minority applicants, they can be useful in determining what incentives lie at the root of any discrimination that is uncovered. In other words, default rates may provide a means of distinguishing between statistical discrimination and bigotry.

It is worth asking why we are interested in identifying the source of mortgage market discrimination. After all, we have argued that this distinction is not important under the law. Nor is the root cause of discrimination likely to interest an individual victim. From a policy-

■ **17** Arrow (1972a, 1972b) and Phelps (1972) offer the first discussion of statistical discrimination; they do so in the context of labor markets.

■ **18** Such variables must pass the "disparate impact test," as we discuss below. See the Federal Financial Institution Examination Council's 1994 joint statement on fair-lending enforcement.

maker's perspective, however, it is quite important to know whether discrimination is belief-based or preference-based. Any policy's ability to eradicate discrimination will depend on the underlying source of that discrimination.<sup>19</sup>

For example, an appropriately designed penalty/subsidy scheme might counteract a bigoted lender's willingness to forgo profitable opportunities for lending to minority applicants, thereby inducing it to equalize its required  $q_A^*$  and  $q_B^*$ .<sup>20</sup> Alternatively (and perhaps more effectively), policymakers may attempt to combat taste-based discrimination by promoting competition in the mortgage market.

While these kinds of policies may work to eliminate bigotry, they can have the opposite effect on a statistical discriminator. Eradication of belief-based discrimination hinges on a policy's ability to make profit-maximizing lenders ignore costless information (the correlation between race and default) that, if used, would improve their profitability. Increased competition strengthens a lender's desire to employ such information, effectively increasing its incentives to statistically discriminate. Instead, statistical discrimination is best fought by focusing directly on the underlying source of differences across groups.

Likewise, penalty/subsidy schemes (much like the current use of the CRA as a tool for fair-lending enforcement) may be less desirable when the true source of discrimination is statistical correlation between race and creditworthiness. Although such a scheme may well enhance minorities' access to credit, it will do so at a very large cost: By making mortgage loans more attractive to less-creditworthy minority applicants, a penalty/subsidy solution would reduce the average creditworthiness of the applicant pool, *strengthening* the correlation between race and default. As a result, even larger subsidies or penalties would be required to eliminate statistical discrimination.

Finally, because statistical discrimination is motivated by a desire to maximize profits, in order to equalize  $s_A^*$  and  $s_B^*$  banks will have to hold whites to a higher creditworthiness standard than minorities ( $q_A^* > q_B^*$ ). That is, to eradicate statistical discrimination, regulators must give banks a "taste for discrimination" against white applicants so that banks will have the incentive to pass up more profitable loans (to whites) in order to make less profitable ones

(to minorities). Beyond its social and political implications, such a policy may be particularly difficult to implement.

Despite its importance from a policy perspective, uncovering the root cause of discrimination has not been the purpose of existing research. The reason is that both the statistical discriminator and the bigot will require a minority applicant to have a "better" application than a white applicant in order to be granted a loan. That is, a finding that  $s_A^* < s_B^*$  is equally consistent with both preference-based and statistical discrimination.

Instead, the difference between taste- and preference-based discrimination is the fact that a bigot will require minorities to meet a higher minimum (inferred) standard of creditworthiness than whites (that is,  $q_A^* < q_B^*$ ), while the statistical discriminator, because he has no taste for discrimination, is unwilling to forgo lending to creditworthy minority applicants and so holds minorities and whites to the same minimum credit standard (that is,  $q_A^* = q_B^*$ ).

This fact suggests that conditional and marginal default rates may help determine the underlying source of any discrimination in the mortgage market. In particular, if conditional default rates are equal across groups, then statistical discrimination is an unlikely explanation for any discrimination that is uncovered through denial-rate analyses. On the other hand, if the conditional default rate of marginal minority borrowers exceeds that of marginal white borrowers, then bigotry becomes a less likely candidate for the root cause of discrimination.

Alternatively, Craig et al. (1998) propose a method for distinguishing empirically between statistical and preference-based discrimination that uses denial-rate data like those employed by Munnell et al. Unlike Munnell et al., they can address this issue because they take advantage of the structure inherent in the underwriting model developed by Longhofer and Peters (1999) and discussed here. Because the difference between statistical discrimination and bigotry lies in the relationship between  $q_A^*$  and  $q_B^*$ , the empiricist must be able to reconstruct the lender's Bayesian updating process (the underwriting process) to arrive at each applicant's inferred quality,  $q(s)$ . A finding that minority applicants are more likely to be rejected than whites, *conditional on*  $q(s)$ , would be consistent with preference-based discrimination. On the other hand, if the likelihood of being rejected varies across racial groups controlling for  $s$ , but does not vary after controlling for each applicant's  $q$ , statistical discrimination is the more plausible explanation.

■ 19 See Longhofer (1995) for a related discussion.

■ 20 Whether this involves the bank's raising its standard for whites or lowering it for minorities, however, is a more complicated question.

## IV. Conclusion

We have shown here how the current mortgage discrimination debate has suffered because of inadequate theoretical underpinnings. Using a formal model of the underwriting process developed by Longhofer and Peters (1999), we are able to define what is meant by discrimination and to design tests for uncovering such behavior. Furthermore, we are able to define several different types of default rates precisely and to explore their implications under different underlying discriminatory incentives. Finally, we argue that developing appropriate policy responses to mortgage market discrimination depends crucially on understanding its root causes. Using our theoretical model, we are able to design some tests for uncovering this crucial information.

The definition of discrimination that arises out of our model implies that denial rate analyses like those performed by Munnell et al. are the proper tool for uncovering discrimination. This should not be interpreted, however, as suggesting that their results prove the existence of widespread discrimination in mortgage markets. Compelling as they are, we share some of the well-documented concerns about the veracity and interpretation of their results.<sup>21</sup> In the end, more research will be required to confirm or refute the existence of widespread discrimination and to understand its causes.

## References

- Arrow, Kenneth J.** "Models of Job Discrimination," in Anthony H. Pascal, ed., *Racial Discrimination in Economic Life*. Lexington, Mass.: Lexington Books, 1972a, pp. 83–102.
- . "Some Mathematical Models of Race Discrimination in the Labor Market," in Anthony H. Pascal, ed., *op. cit.*, 1972b, pp. 187–203.
- Becker, Gary S.** *The Economics of Discrimination*, 2d. ed. Chicago: University of Chicago Press, 1971.
- . "The Evidence against Banks Doesn't Prove Bias," *Business Week*, April 19, 1993a, p. 18.
- . "Nobel Lecture: The Economic Way of Looking at Behavior," *Journal of Political Economy*, vol. 101, no. 3 (June 1993b), pp. 385–409.
- Berkovec, James A., Glenn B. Canner, Stuart A. Gabriel, and Timothy H. Hannan.** "Race, Redlining, and Residential Mortgage Loan Performance," *Journal of Real Estate, Finance, and Economics*, vol. 9, no. 3 (November 1994), pp. 263–94.
- Brimelow, Peter, and Leslie Spencer.** "The Hidden Clue," *Forbes*, January 4, 1993, p. 48.
- Browne, Lynn Elaine, and Geoffrey M.B. Tootell.** "Mortgage Lending in Boston—A Response to the Critics," Federal Reserve Bank of Boston, *New England Economic Review*, September/October 1995, pp. 53–78.
- Calem, Paul, and Michael Stutzer.** "The Simple Analytics of Observed Discrimination in Credit Markets," *Journal of Financial Intermediation*, vol. 4, no. 3 (July 1995), pp. 189–212.
- Calomiris, Charles W., Charles M. Kahn, and Stanley D. Longhofer.** "Housing-Finance Intervention and Private Incentives: Helping Minorities and the Poor," *Journal of Money, Credit, and Banking*, vol. 26, no. 3, pt. 2 (August 1994), pp. 634–74.
- Craig, Ben R., Stanley D. Longhofer, and Stephen R. Peters.** "Measuring Creditworthiness and Discrimination in Mortgage Lending," Federal Reserve Bank of Cleveland, unpublished manuscript, May 1998.

■ 21 See, for example, Horne (1994, 1997), Yezer et al. (1994), Hunter and Walker (1996), and Day and Leibowitz (1998). See also Browne and Tootell (1995).

- Day, Theodore E., and Stanley J. Leibowitz.** "Mortgage Lending to Minorities: Where's the Bias?" *Economic Inquiry*, vol. 36, no. 1 (January 1998), pp. 3–28.
- Federal Financial Institutions Examination Council.** "Policy Statement on Discrimination in Lending," pt. III, *Federal Register*, vol. 59, no. 73 (April 15, 1994).
- Ferguson, Michael F., and Stephen R. Peters.** "What Constitutes Evidence of Discrimination in Lending?" *Journal of Finance*, vol. 50, no. 2 (June 1995), pp. 739–48.
- , and —. "Cultural Affinity and Lending Discrimination: The Impact of Underwriting Errors and Credit Risk Distribution on Applicant Denial Rates," *Journal of Financial Services Research*, vol. 11, nos. 1–2 (April 1997), pp. 153–67.
- , and —. "Is Lending Discrimination Always Costly?" University of Illinois, unpublished manuscript, September 1998.
- Galster, George C.** "The Facts of Lending Discrimination Cannot Be Argued Away by Examining Default Rates," *Housing Policy Debate*, vol. 4, no. 1 (1993), pp. 141–46.
- Horne, David K.** "Evaluating the Role of Race in Mortgage Lending," *FDIC Banking Review*, vol. 7, no. 1 (Spring/Summer 1994), pp. 1–15.
- . "Mortgage Lending, Race, and Model Specification." *Journal of Financial Services Research*, vol. 11, nos. 1–2 (April 1997), pp. 43–68.
- Hunter, William C., and Mary Beth Walker.** "The Cultural Affinity Hypothesis and Mortgage Lending Decisions," *Journal of Real Estate Finance and Economics*, vol. 13, no. 1 (July 1996), pp. 57–70.
- Longhofer, Stanley D.** "Rooting Out Discrimination in Home Mortgage Lending." Federal Reserve Bank of Cleveland, *Economic Commentary*, November 1995.
- , and **Stephen R. Peters.** "Self-Selection and Discrimination in Credit Markets," Federal Reserve Bank of Cleveland, Working Paper No. 9809, July 1998 (revised February 1999).
- Macey, Jonathan R.** "Banking by Quota," *Wall Street Journal*, September 7, 1994, p. A14.
- Munnell, Alicia H., Lynn E. Browne, James McEneaney, and Geoffrey M.B. Tootell.** "Mortgage Lending in Boston: Interpreting HMDA Data," Federal Reserve Bank of Boston, Working Paper No. 92–7, October 1992.
- , **Geoffrey M.B. Tootell, Lynn E. Browne, and James McEneaney.** "Mortgage Lending in Boston: Interpreting HMDA Data," *American Economic Review*, vol. 86, no. 1, March 1996, pp. 25–53.
- Phelps, Edmund S.** "The Statistical Theory of Racism and Sexism," *American Economic Review*, vol. 62, no. 4 (September 1972), pp. 659–61.
- Tootell, Geoffrey M.B.** "Defaults, Denials, and Discrimination in Mortgage Lending," *New England Economic Review*, (September/October 1993), pp. 45–51.
- Yezer, Anthony M.J., Robert F. Phillips, and Robert P. Trost.** "Bias in Estimates of Discrimination and Default in Mortgage Lending: The Effects of Simultaneity and Self-Selection," *Journal of Real Estate Finance and Economics*, vol. 9, no. 3 (November 1994), pp. 197–215.

# Banking and Commerce: How Does the United States Compare to Other Countries?

by João A.C. Santos

João A.C. Santos is an economist at the Bank for International Settlements in Basel, Switzerland. The author thanks Joseph Haubrich and George Pennacchi for their useful comments and suggestions. Part of this research was developed while the author was at the Federal Reserve Bank of Cleveland.

## Introduction

Recent attempts to repeal the Glass-Steagall Act, which would allow commercial banking to mix with investment banking, have again initiated debate over current regulation prohibiting the affiliation between banking and commerce. Although participants in that debate have put forward many arguments based on U.S. historical experience, international evidence, and research results, the debate remains unsettled.

There is no consensus as to the extent of the affiliation between banks and nonfinancial firms throughout American history. In some cases, this lack of consensus is the result of differing definitions of banking and commerce. In other cases, it results from considering different ways through which the corporate affiliation between banks and firms can be implemented. Still in other cases, the difference results from the use of like terminology, “mixing banking with commerce,” to indicate two different ways to implement the affiliation—for a bank to own a stake in a firm, or for a firm to own a stake in a bank. As we will see, each of these two forms of affiliation has been subject to different regulations. Therefore, using either one of them as a basis for measuring the association between banking and commerce will lead to different conclusions.

Any dividing line between banking and commerce is bound to be imprecise. The legal definition of banking itself has changed over time and has been subject to many interpretations. Throughout this article, the term *banking* is used to denote commercial banking in its simplest form; therefore, a bank is defined as a firm that accepts demand deposits and makes loans. This allows, when necessary, a distinction between commercial banking and investment banking, which includes activities such as brokerage and securities underwriting. *Commerce* is used to denote all nonfinancial firms—that is, commercial and industrial firms. The article assumes that the association between banks and firms may be implemented (1) through one party undertaking the other party’s activity in-house; (2) through one party’s investment in the capital of the other party; or (3) through a parent company, such as a bank holding company (BHC), which owns a stake in both a bank and a firm.<sup>1</sup>

■ 1 The case where an investor, rather than a corporation, owns a stake in a bank and in a firm is also sometimes considered as a way to mix banking with commerce. This form of association at the personal level is not considered here. See Huertas (1986) for examples of investors in the United States with simultaneous controlling interests in banks and in nonfinancial firms.

Using these definitions, the article shows that, traditionally, U.S. banking regulations have not allowed banks to enter commerce, though there have been significant exceptions to this separation at different points in time. Regulation of firms' ownership of banks is very recent, beginning only with the passage of the Bank Holding Company Act (BHCA) in 1956. When comparing current U.S. regulations on the affiliation between banks and firms to those in force abroad, the article finds that other countries' regulations are significantly more liberal. However, an analysis of banks' investments in equities in these countries shows that such investments account for a small fraction of banks' total assets.

As in the past, the most recent debate on whether banks should be permitted to affiliate with firms has focused on the potential implications of that affiliation for the banking sector and the safety net.<sup>2</sup> There has also been a great deal of interest in the discussion of how these implications would vary with the organizational forms adopted by banks to implement the affiliation.<sup>3</sup> The potential effects of that affiliation on other segments of the economy, such as nonfinancial firms, have been largely ignored, yet they are of considerable importance. Despite recent research on the design of financial systems, our understanding of the implications associated with different financial systems is far from complete. Nonetheless, as we will see, there is a segment of that research that is particularly timely for the ongoing debate. This research has focused on the implications for borrowing firms resulting from banks' owning equity positions in these firms.

The paper proceeds as follows: The next section summarizes the regulations on banks' investments in firms and those on firms' investments in banks throughout American history. Section II compares the current U.S. regulations with those in force abroad and discusses the extent of banks' investments in equities in several countries. Section III reviews the literature dealing with the implications of banks' investments in nonfinancial firms. Section IV ends the paper with some final remarks.

■ **2** See Huertas (1988), Mester (1992), or Saunders (1994) for a review of the arguments at the center of the debate.

■ **3** Some of the issues in that discussion have also been present in the debate over the organizational forms that banks should be allowed to adopt to combine commercial banking with investment banking. See Santos (1998a) for an extensive discussion of these issues in the context of the commercial–investment banking debate and Cumming and Sweet (1987) for a comparison of the predominant organizational forms adopted in the G-10 countries to integrate banking with commerce.

## I. Banking and Commerce Throughout American History

Although the debate on the association between banks and nonfinancial firms prior to the National Banking Act of 1864 still continues, since then, and in particular after the passage of the BHCA in 1956, the law has dictated the separation of banking from commerce.<sup>4</sup> In the past, the regulation of firms' ownership of stakes in banks was not as restrictive as that of banks' ownership of stakes in firms. With respect to the current U.S. banking regulations, they allow firms to make limited investments in banks, but not to control them. That regulation also prohibits banks from making investments in firms, but allows BHCs to make such investments within certain limits.

### Nonfinancial Firms' Investment in Banks

Restrictions on nonfinancial firms' investment in banks first appeared in American banking regulation in 1956 with the BHCA. Until then, any commercial or industrial firm could be the sole owner of a bank. The primary restrictions on the ownership of banks before the BHCA were not directed at firms. One of these restrictions was defined in the National Banking Act of 1864 and prohibited a bank from owning shares in another bank. Another restriction was defined in section 20 of the Glass-Steagall Act and prohibited any firm "principally engaged" in the investment banking business from affiliating with member banks.<sup>5</sup>

There are many examples of firms owning banks throughout American history. Firms' ownership of banks goes as far back as 1799, when the Manhattan Company was chartered to supply New York City with fresh water. That company entered the business of banking by creating the Bank of the Manhattan Company, which it operated as a subsidiary. Likewise, in 1954 the Transamerica Corporation controlled banks in five western states and owned subsidiaries engaged in various nonbanking activities.<sup>6</sup>

■ **4** See Shull (1983), Huertas (1986), Fein and Faber (1986), and Blair (1994) for an analysis of the association between banking and commerce in American banking history.

■ **5** An additional constraint, though not very restrictive, of the Banking Act of 1933 was to require corporations owning more than 50 percent of one or more member banks to apply to the Federal Reserve for a permit to vote their stock.

With respect to the current restrictions on firms' investments in banks, they are determined by three definitions contained in the BHCA of 1956 and its subsequent 1966 and 1970 amendments. They define a BHC, a bank, and the powers of a BHC, respectively. When the BHCA was enacted in 1956, it defined a BHC as any company that controlled two or more banks. Therefore, single BHCs were exempted. Not surprisingly, many single BHCs were established and their number continued to grow until 1970, when the BHCA was amended to end the loophole. The BHCA states that a company controls another company if it (1) has direct or indirect ownership, control, or power to vote 25 percent or more of any class of voting securities of that company; (2) is able in any manner to elect a majority of the directors of that company; or (3) is able to exercise, directly or indirectly, a controlling influence over that company as determined by the Board. That definition, therefore, allows firms to make noncontrolling investments of up to 25 percent of the voting shares of a bank.

As expected, the definition of a bank has also played a key role in the regulations that separate banking from commerce. As mentioned above, the Glass-Steagall Act prohibition against firms "principally engaged" in investment banking owning banks applied only to member banks—that is, all national banks and the state banks that chose to be members of the Federal Reserve System. The BHCA of 1956, however, adopted a broader definition of a bank: "[A]ny national banking association or any state bank, savings bank or trust company" was considered a bank. The 1966 amendment to the BHCA redefined a bank as any institution that accepted demand deposits. The 1970 amendment again redefined banks, this time as any firm that "(1) accepts deposits that the depositor has a legal right to withdraw on demand, and (2) engages in the business of making commercial loans." As a result of this last definition, any institution that offered one service but not the other would not be classified as a bank.<sup>7</sup> Such firms became known as "nonbank banks" and were owned by many different corporations. This led to yet another redefinition of banks in the Competitive Equality Banking Act of 1987 as an institution that is either insured by the Federal Deposit Insurance Corporation or offers demand deposits (or transaction accounts) and makes commercial loans.<sup>8</sup>

Finally, the BHCA restricted BHCs' powers to the business of banking and to some banking-related activities. The act contained a list of several nonbanking businesses that BHCs could

offer and directed the Board to authorize any other nonbanking activities. The 1970 amendment, however, requires these nonbanking activities to be "closely related to banking" and for their provision by BHCs or their subsidiaries to produce public benefits that outweigh possible adverse effects.<sup>9</sup>

In sum, a nonfinancial firm in the United States may make equity investments in banks and BHCs, but they cannot control these institutions. More specifically, nonfinancial firms cannot own more than 25 percent of a bank or a BHC because they would themselves become BHCs. As for the BHCs, they can be involved in banking and banking-related business, but not in nonfinancial activities.

### Banks' Investment in Nonfinancial Firms

Throughout American banking history, banks were not permitted to own nonfinancial firms. There were, however, notable exceptions to this rule. Shull (1983) argues that the U.S. policy of prohibiting banks from owning firms originated in the 1694 House of Commons Act establishing the Bank of England. That act defined the powers of the Bank of England and included a clause explicitly prohibiting the Bank from dealing in merchandise.

The charters of early U.S. banks did not always define the business of banking or bank powers. There were, however, banks whose charters included clauses mimicking that of the Bank of England. Examples of such banks include the Bank of North America, established in 1781 as the first incorporated bank in the United States; the Bank of New York, established in 1784; and the First and Second Banks

■ **6** See Huertas (1986) and Shull (1994) for more examples of firms that owned banks at various times in the nineteenth and twentieth centuries.

■ **7** Some institutions, for example, continued to offer demand deposits, but restricted their extension of commercial credit to the purchase of money-market instruments such as commercial paper. Other firms avoided that definition of banks by offering NOW (negotiable orders of withdraw) accounts instead of demand deposits. NOWs are similar to demand deposits, with the difference that they require prior notice before the customer can withdraw the money.

■ **8** The nonbank banks established at the time were grandfathered, with the restriction that their assets could not grow more than 7 percent in any 12-month period. See Mester (1992) and Shull (1994) for several examples of nonbank banks.

■ **9** See Pollard, Passaic, Ellis, and Daly (1988) for a detailed presentation of U.S. banking law and the list of nonbanking activities allowed by the Board.

of the United States, established in 1791 and 1816, respectively.

With time, bank charters increasingly detailed the powers of banks. One of the first definitions of the “business of banking” appeared in 1825 in the charter of the Commercial Bank of Albany. It included a detailed list of the bank’s powers and a clause requiring the bank to have “no other powers whatsoever,” implicitly prohibiting the bank from dealing in merchandise. This definition was later included in the New York Free Banking Act of 1838, which opened banking to the general public, and it influenced the banks’ powers clause of the National Banking Act of 1864, which established the national banking system.

The National Banking Act allowed national banks to offer “all such incidental powers as shall be necessary to exercise the business of banking.” The act listed the explicit powers available to banks and gave the Comptroller of the Currency the authority to determine the activities that were incidental to the business of banking. Since its enactment, the act’s powers clause has been the subject of much debate and interpretation. In litigation, it was determined that national banks could accept corporate stock as collateral or payment for debt (in this case, they could hold the stock only for a limited period of time), but they could not deal in or purchase stock as an investment.<sup>10</sup> In addition, banks could not engage in the operation of a business, even if it had been acquired in satisfaction of a debt. Therefore, the National Banking Act extended in time the prohibition of the early bank charters against dealing in merchandise.

Throughout American history there were, however, periods of time when banks were permitted to enter commerce. For example, there is evidence that during the second quarter of the nineteenth century—a period that became known as the “free” or “wildcat” banking era—banks in some states, such as Connecticut, Michigan, New Jersey, South Carolina and Texas, received charters allowing them to combine banking with many other nonbanking activities. There is also evidence of banks mixing with commerce after the enactment of the National Banking Act in 1864. When the act was enacted, it was believed that state-chartered banks would convert to national charters. But while a national charter did confer greater prestige and competitive advantage in terms of note issuance (state-chartered banks had to pay taxes on the notes they issued), state charters allowed broader powers. California, for example, allowed its banks to enter commerce.

Another instance in which banking mixed with commerce occurred during the post-Civil War period through private banks. These were unincorporated banks that did not issue notes and were free to pursue any activity. Although they were referred to as banks, many states prohibited them from using the name “bank.” Private banks combined commercial banking with many other activities, including brokerage, securities underwriting, and commerce. With the passage of the Glass-Steagall Act in 1933, some private banks opted to terminate their deposit-taking activities in order to continue their nonbanking activities. Some of them went on to become leaders in the investment banking business in the United States.<sup>11</sup>

Current regulation prohibits banks from making investments in nonfinancial firms. As mentioned above, the National Banking Act of 1864 has been interpreted as prohibiting national banks from making such investments. This prohibition has also been extended to state member banks. State nonmember banks are usually limited to investments that are permissible for national banks, but there are exceptions which vary with the state in which they are chartered.<sup>12</sup> Finally, with respect to BHCs, as noted before, the BHCA restricts them to the business of banking and to some banking-related activities. Because of this, BHCs have only been allowed to make investments in nonfinancial firms that do not account for more than 5 percent of the firm’s outstanding voting shares.<sup>13</sup>

In conclusion, banks in the United States in general have not been allowed to make investments in the capital of nonfinancial firms. There

■ 10 See James (1995) for a characterization of the conditions under which banks can accept corporate stock for payment of debt claims when the borrower is in financial distress.

■ 11 See Fein and Faber (1986) for examples of banks that combined banking with nonbanking activities during the free banking era, and for examples of private banks that mixed banking with nonbanking businesses.

■ 12 See Halpert (1988) for some examples of these exceptions.

■ 13 In addition to the restrictions emanating from the banking law, it is usually argued that the legal doctrine of equitable subordination reduces banking organizations’ incentive to make equity investments in firms. This doctrine, which is in force in the United States and the United Kingdom, reduces a bank’s incentive to take an equity position in a firm to which it has extended loans, because the exercise of control rights associated with its equity stake may lead to a loss of the bank’s legal status as a creditor in the event of bankruptcy. See Prowse (1990) and Roe (1990) for a characterization of the doctrine and Kroszner (1998) for a discussion of its implications for banking organizations.

TABLE 1

### Regulations on Nonfinancial Firms' Ownership of Commercial Banks

Austria	Complies with EC Second Banking Directive. <sup>a</sup>	Japan	Total investment is limited to firms' capital or net assets. The Anti-Monopoly Law prohibits establishment of a holding company whose main business is to control the business activities of other domestic companies through the holding of ownership.
Belgium	Complies with the EC Second Banking Directive. <sup>a</sup> However, the Banking and Finance Commission examines the "fit and proper" character of those shareholders holding at least 5 percent of the bank's capital.	Luxembourg	Nonfinancial firms may legally be the majority shareholders in banks. However, general policy discourages nonfinancial groups or private persons from being major shareholders in banks.
Canada	No shareholder may own more than 10 percent of a bank's outstanding shares.	Netherlands	Complies with the EC Second Banking Directive. <sup>a</sup> However, a declaration of nonobjection from the Minister of Finance is required for an investment exceeding 5 percent of a bank's capital.
Denmark	Complies with the EC Second Banking Directive. <sup>a</sup> However, a firm may not have an interest that allows a decisive influence on the bank.	Portugal	Complies with the EC Second Banking Directive. <sup>a</sup>
Finland	Complies with the EC Second Banking Directive. <sup>a</sup> However, the firm cannot vote at an annual meeting with more than 5 percent of the total voting rights present at the meeting.	Spain	Complies with the EC Second Banking Directive. <sup>a</sup> However, a nonfinancial firm cannot hold more than 20 percent of the shares of a new bank during the first five years of its existence. Specified shareholder thresholds require authorization by the Bank of Spain before additional investment.
France	Complies with the EC Second Banking Directive. <sup>a</sup>	Sweden	Ownership is limited to 50 percent except when a bank is near insolvency and there is a need for external capital injection. In this case, greater ownership may be permitted based on the suitability of new owners.
Germany	Complies with the EC Second Banking Directive. <sup>a</sup>	Switzerland	Unrestricted.
Greece	Complies with the EC Second Banking Directive. <sup>a</sup>	United Kingdom	Complies with EC Second Banking Directive. <sup>a</sup>
Ireland	Advance notification is required for any application of more than 5 percent of the voting rights in a bank, and prior approval is required for any application of 10 percent or more of the total shares or voting rights or any holding or interest that confers a right to appoint or remove directors.		
Italy	Persons who engage in significant business activity in sectors other than banking and finance are forbidden from acquiring an equity stake which, when added to that already held, would result in a holding exceeding 15 percent of the voting capital of a bank or in the control of a bank.		

a. The EC Second Banking Directive subjects a "qualifying investment" (a direct or indirect holding in an undertaking equal to at least 10 percent of its capital or voting rights or permitting the exercise of significant influence over its management) to regulatory consent based only on the suitability of shareholders.

SOURCE: Barth, Nolle, and Rice (1997).

were, however, periods of time when certain banks were granted permission to do so and other occasions when they were able to combine banking with commerce by developing substitutes for banks. As for the current regulation, it prohibits banks from investing in nonfinancial firms and it only allows BHCs to make some rather limited investments in these firms.

## II. Banking and Commerce: International Evidence

Not all countries regulate the association between banking and commerce as restrictively as the United States. The difference is particularly striking in the case of the regulations on banks' investment in equities of nonfinancial firms. As this section will show, most of the countries reviewed here allow banking and commerce to mix. However, the data available indicate that, on an aggregate basis, shares and participations represent a small fraction of foreign banks' assets.

TABLE 2

### Regulations on Commercial Banks' Ownership of Nonfinancial Firms

Austria	Complies with EC Second Banking Directive. <sup>a</sup>	Luxembourg	Complies with EC Second Banking Directive. <sup>a</sup>
Belgium	Single shareholdings may not exceed 10 percent of a bank's own funds, and such shareholdings on an aggregate basis may not exceed 35 percent of own funds. More restrictive than the EC Second Banking Directive during a transitional period. <sup>a</sup>	Netherlands	Complies with EC Second Banking Directive. <sup>a</sup> However, a declaration of nonobjection from the Minister of Finance is required for any bank investment exceeding 10 percent of the capital of the firm.
Canada	Limited to 10 percent of outstanding shares of a nonfinancial firm, with aggregate holdings not to exceed 70 percent of bank capital.	Portugal	Complies with EC Second Banking Directive. <sup>a</sup> However, a bank may not control more than 25 percent of the voting rights of a nonfinancial firm.
Denmark	Complies with EC Second Banking Directive. <sup>a</sup> However, a bank may not hold a permanent, decisive participation in a nonfinancial firm.	Spain	Complies with EC Second Banking Directive. <sup>a</sup>
Finland	Complies with EC Second Banking Directive. <sup>a</sup>	Sweden	Investments on an aggregated basis are limited to 40 percent of a bank's own funds. Ownership in a firm is limited to 5 percent of this base and must not exceed 5 percent of the total voting power of the firm. These limits do not apply when a bank must protect itself against credit losses; in this case, the bank must sell when market conditions are appropriate.
France	Complies with EC Second Banking Directive. <sup>a</sup>	Switzerland	A single participation is limited to the equivalent of 20 percent of the bank's capital; however, the Swiss Banking Commission can allow this limit to be exceeded.
Germany	Complies with EC Second Banking Directive. <sup>a</sup>	United Kingdom	Complies with EC Second Banking Directive. <sup>a</sup> However, an ownership share of more than 20 percent requires the investment to be deducted from the bank's capital when calculating its capital adequacy on a risk basis. Otherwise, the investment is treated as a commercial loan for the risk-based calculation.
Greece	Complies with EC Second Banking Directive. <sup>a</sup>		
Ireland	Complies with EC Second Banking Directive. <sup>a</sup>		
Italy	Most banks are subject to an overall investment limit of 15 percent of their own funds (7.5 percent in the case of unlisted firms) and to a concentration limit of 3 percent of funds in each holding in nonfinancial firms. Some banks, due to their size and proven stability, are subject to less stringent limits (overall and concentration limits of 50 percent and 6 percent, respectively, for leading banks, and 60 percent and 15 percent for specialized banks).		
Japan	A single bank's ownership is limited to 5 percent of a single firm's shares.		

a. The EC Second Banking Directive limits qualifying investments to no more than 15 percent of a bank's own funds for investment in a single firm, and to no more than 60 percent for all investment in nonfinancial firms. In exceptional circumstances, these limits may be exceeded, but the amount by which the limits are exceeded must be covered by a bank's own funds and these funds may not be included in the solvency-ratio calculation.

SOURCE: Barth, Nolle, and Rice (1997).

### The Regulations

Tables 1 and 2 summarize, for a selected group of foreign countries, the main regulations on firms' ownership of banks and on banks' ownership of firms, respectively. An analysis of these tables indicates that, as in the United States, the majority of countries regulate firms' ownership of banks less restrictively than they do banks' ownership of firms. With respect to firms' investment in banks, most countries' regulations do not impose an absolute limit on the share of a bank's capital that a firm can own. Instead, their regulations give supervisory authorities control over the suitability of banks' shareholders. There are, however, a few countries, such as Canada, Italy, Sweden, and the United States, that limit the share of a bank's capital that a firm can own.

Contrary to the United States, the foreign countries reviewed allow for limited investments in firms by banks. Most regulations define a limit for each individual investment and for the total of these investments in terms of the bank's capital. Some countries put an additional restriction on that investment defined in terms of the firm's capital. For example, in the Netherlands banks need authorization to own a stake larger than 10 percent. In the United Kingdom, stakes larger than 20 percent are deducted from the bank's capital for the purpose of capital adequacy. In Denmark, banks are not permitted to hold a permanent participation in firms. Finally, in Portugal, Canada, and Japan, banks may not own more than 25 percent, 10 percent, and 5 percent of the capital of a nonfinancial firm, respectively.

TABLE 3

Banks' Investment in Equities<sup>a</sup>

Country	N <sup>b</sup>	Loans/ assets	Securities/ assets	Ratio of shares and participations to:		
				Assets	Nonbank deposits <sup>c</sup>	Capital reserves
Austria	1,041 <sup>d</sup>	0.509	0.143	0.038	0.087	0.882
Belgium	104	0.362	0.330	0.017	0.47	0.604
Large banks	7	0.376	0.307	0.021	0.055	0.787
Canada	11	0.665	0.196	—	—	—
Denmark	114 <sup>e</sup>	0.433	0.290	0.042	0.075	0.604
Finland	7	0.484	0.291	0.052	0.098	1.142
France	413 <sup>f</sup>	0.340	0.188	0.028	0.116	0.844
Germany	266	0.574	0.170	0.048	0.116	0.902
Large banks	3	0.541	0.185	0.063	0.132	1.234
Greece	18	0.270	0.336	0.043	0.060	0.916
Large banks	4	0.249	0.367	0.049	0.068	1.134
Italy	269 <sup>g</sup>	0.424	0.139	0.020	0.055	0.220
Japan <sup>h</sup>	139	0.668	0.143	0.046	0.060	1.387
Large banks	11	0.650	0.126	0.060	0.083	1.949
Luxembourg	220 <sup>i</sup>	0.189	0.189	0.003	0.008	0.119
Netherlands	173 <sup>j</sup>	0.634	0.141	0.006	0.014	0.153
Portugal	37	0.333	0.232	0.031	0.060	0.382
Spain	170 <sup>k</sup>	0.411	0.184	0.041	0.091	0.477
Sweden	13	0.436	0.356	0.029	0.056	0.494
Switzerland	88	0.474	0.178	0.049	0.106	0.766
Large banks	4	0.471	0.181	0.052	0.113	0.860
U.K.	40	0.521	0.185	—	—	—
U.S.	9,986	0.634	0.214	—	—	—
Large banks	100	0.625	0.187	—	—	—

a. Data relate to December 31, 1995.

b. Number of commercial banks, except when indicated otherwise.

c. Excludes interbank deposits.

d. Includes commercial, foreign-owned, savings, and cooperative banks and other financial institutions.

e. Includes commercial and savings banks.

f. Includes foreign-controlled banks operating in France in the form of branches or subsidiaries. However, branches of banks with headquarters in other EC countries are excluded.

g. Includes limited company, cooperative, and main mutual banks, central credit institutions, and branches of foreign banks.

h. Data relate to fiscal year ending March 31, 1995.

i. Includes banks and *sociétés anonymes* set up under Luxembourg law and foreign companies.

j. Includes universal banks, banks organized on a cooperative basis, savings banks, mortgage banks, other capital market institutions, and security credit institutions.

k. Includes foreign-owned banks.

SOURCE: Organisation for Economic Co-operation and Development.

Banks' Investment  
in Equities

The firm-level data necessary to study the association between banking and commerce across countries are not readily available. The OECD, however, publishes data that give us an idea of the extent of banks' investment in equities.

Table 3 presents some statistics computed with data for the same countries, with the exception of Ireland, that were included in tables 1 and 2. When reading table 3, the reader should keep in mind that the statistics were computed using aggregated data and that the variable "shares and participations" include the investments in equities of all corporations, not only nonfinancial firms. In addition, the reader should also take into account the usual caveats, such as the differences in accounting rules and reporting methods, when making international comparisons.

As table 3 shows, for most countries, loans represent a larger portion of the banks' assets than securities (Greece and Luxembourg are the only exceptions), and shares and participations represent a small fraction of banks' securities. Banks' limited investments in equities are also evidenced by two other statistics: banks' investments in shares and participations represent less than 12 percent of their nonbanks' deposits and are smaller than banks' capital and reserves (Finland and Japan are the only two exceptions).<sup>14</sup> Table 3, however, tells us nothing about the extent of that investment at the bank level. The data available allow us to study that issue only for a given group of banks, the group of the largest banks, and only for a subset of the countries considered here (Belgium, Germany, Greece, Japan, Switzerland, and the United States). Comparing the statistics for these banks with those for the banking sector in the same country, it becomes clear that larger banks are more involved in equities investment than the rest of the banking sector in the country, though the extent of their investments is still quite limited. For example, in Germany and Japan, the two countries where the largest banks' investment in equities is highest, shares and participations still represent less than 7 percent of the banks' assets.

Table 3 is also mute with respect to several other important issues related to banks' investments in equities. For example, on average, those investments account for what stake of the

■ 14 The banking sector's limited investment in equities in 1995 evidenced in table 3 accords with the results unveiled in Langohr and Santomer (1985) based on comparable data for 1981.

firms' capital? At what point in a firm's life do banks make these investments? Are they made in conjunction with other services, such as the extension of a loan or a line of credit? As we will see in the next section, some of these questions have already been the subject of research, mainly using data on Japan and Germany.

### III. Banks' Investment in Equities: A Review of the Literature

As with previous debates on changes to the U.S. banking regulation, the ongoing debate over a relaxation of the barriers between banking and commerce has focused on the potential implications of this regulatory change on banks' stability. The consequences of allowing banks to mix with nonfinancial firms, however, go far beyond that impact. For example, commercial banks' ability to make investments in nonfinancial firms would allow them to own stakes in firms to which they extend loans. Such investments would influence, among other things, banks' relationships with their borrowers, including the way they monitor borrowers and the design of contracts that banks could offer borrowers throughout their life. These effects would, in turn, influence borrowers' incentives, as well as the cost and amount of funds made available to them.

For many years, economic theory has focused on the real sector of the economy and disregarded the financial sector, viewing it as a veil. The justification was that in a frictionless world, à la Arrow–Debreu, there is no room for financial intermediaries. However, the world we live in is quite different from that envisioned by Arrow and Debreu, and financial intermediaries are perceived to play a key role in it. For example, Gerschenkron (1962) argues that financial intermediation influences long-term growth. Mayer (1988) suggests that there are systematic differences in performance between the so-called bank-based systems and the market-based systems, in which banks play a lesser role and financial markets are more prominent. King and Levine (1993) find a strong correlation between the size of the financial system and the level of economic development.

Research on financial system design is still in its early stages, but it has already unveiled some important implications of certain features of a financial system.<sup>15</sup> For example,

Dewatripont and Maskin (1990) argue that bad projects persist too long in bank-based systems, whereas good projects are cut off too early in market-based systems. Sabani (1994) suggests that market-based systems restructure less financially distressed borrowers than bank-based systems. Allen and Gale (1997) argue that bank-based systems provide better intertemporal risk sharing, whereas market-based systems provide better cross-sectional risk sharing. Boot and Thakor (1997) justify the coexistence of financial intermediaries and securities markets because each performs a different role: the former resolve postlending and interim moral hazard problems, while the latter facilitate trades by informed agents and, hence, the transmission of information.<sup>16</sup>

Research on financial institution design has sought primarily to explain the existence of banks.<sup>17</sup> More recently, however, that research has focused on the implications of mixing commercial banking with other activities, in particular with investment banking.<sup>18</sup> With respect to the association between banking and commerce, most of it has focused on the potential implications of banks' equity investments on nonfinancial firms.<sup>19</sup> Pozdena (1991), Kim (1992), and John, John, and Saunders (1994), for example, show that a borrower's risk-taking incentive is reduced when the financier funds the borrower through a combination of a loan and an equity stake, rather than through a single loan. Santos (1999) studies the implications of owning an equity stake when funding is provided by a bank rather than a financier in the presence of moral hazard caused by deposit insurance. He shows that allowing banks to make equity investments in firms to which they extend loans does not increase the moral hazard problem. Boyd, Chang, and Smith (1997), however, reach the opposite conclusion. Key to their findings is the assumption that banks can

■ 15 For an extensive discussion of the literature on the design of financial systems, see Thakor (1996) and Allen and Gale (1999).

■ 16 For additional theories explaining the simultaneous existence of securities markets and banks, see Gorton and Haubrich (1987) and Seward (1984).

■ 17 See Bhattacharya and Thakor (1993) and Freixas and Rochet (1997) for an extensive review of the contemporary banking literature.

■ 18 See Santos (1998b) and Rajan (1996) for a review of the literature on the association between commercial and investment banking.

■ 19 See Santos (1998c) for a review of the literature on the affiliation between banking and commerce.

benefit as equity holders of the firm, but not as debt holders, from the borrower's behavior associated with the moral hazard problem embedded in the model. Their model differs from Santos' model in various respects. For instance, they focus on banks' monitoring to control the moral hazard problem, while Santos focuses on the incentives driven by debt and equity contracts.

Rajan (1992) studies the impact of the financier's equity stake on his credibility when he underwrites the securities of firms to which he has loans outstanding. Rajan shows that a possible remedy would be for the financier to commit to purchase an equity stake at market price at the time of the new issue. In contrast, in a model where the financier holds a claim in firms and chooses to underwrite some of them to liquidate his claim, Puri (1996) shows that an equity claim gives him more incentive than a debt claim to underwrite bad firms and retain his equity claim in good ones. Berlin, John, and Saunders (1996) show that having an informed financier holding both a debt claim and an equity stake can prevent him from colluding with the borrower to exploit the firm's nonequity-uninformed stakeholders when the firm is in financial distress.

Haubrich and Santos (1998) focus on a set of issues rather different than those addressed by the previous literature. They use the liquidity approach to financial intermediation pioneered by Myers and Rajan (1998). Myers and Rajan use the positive and negative aspects of liquidity to derive banks: These institutions emerge as a special type of conglomerate, one combining a firm that takes in liquid deposits with a firm making illiquid loans. Haubrich and Santos use a similar analysis to identify the conditions under which a broader class of conglomerates, those combining banks with nonfinancial firms, can be advantageous.

Aside from these theoretical studies, there is already a large body of empirical literature on various aspects of banks' ownership of equity stakes in nonfinancial firms. However, nearly all of that literature relates to German and Japanese banks. Some studies find evidence that bank equity stakes in healthy firms reduce agency costs. Sheard (1989), for example, finds that the bank with the largest loan share is generally one of the top five shareholders of the firm. Prowse (1990) also finds a significant correlation between the percentage of outstanding debt and the percentage of outstanding equity held in the same firm by the largest debt holder; the correlation becomes more significant in firms where

shareholders have greater scope to engage in opportunistic behavior at the expense of debt holders. Flath (1993) also finds that banks in Japan hold more stock in companies for which the agency problems of debt are more severe, and that stockholding induces greater borrowing. As for Germany, Cable (1985) finds that bank voting control and bank representation on a firm's supervisory board are both significantly correlated with bank borrowing by the firm.<sup>20</sup> Chirinko and Elston (1996), however, find that independent firms have more bank debt than bank-influenced firms.<sup>21</sup>

Other studies find evidence that bank equity stakes in healthy firms reduce the incentive and informational problems, thus increasing the availability of funding and reducing the cost of it. Hoshi, Kashyap, and Scharfstein (1991), for example, find that investment by Japanese firms with a close relationship to a bank, in the industrial organizational form of a *keiretsu*, is less sensitive to their liquidity than firms raising funds through more arms-length transactions. Elston (1993) finds similar results for a sample of German firms by comparing investment by firms in which banks have an equity stake, with investment by firms in which banks do not have a direct equity ownership. Weinstein and Yafeh (1998), however, find a less favorable result in Japan. They find that the cost of capital for unaffiliated firms is lower than that for firms with a main bank—that is, firms whose largest lender is also their largest bank equity holder.

Still another set of studies looks at the impact of bank equity stakes on healthy firms' performance. Cable (1985) finds, based on a sample of German firms, that bank voting control and bank representation on a firm's supervisory board are both significantly correlated with bank borrowing by the firm and with the firm's performance. Gorton and Schmid (1995) also find that banks' equity ownership in firms improves the performance of these firms. Chirinko and Elston (1996), however, find that German banks do not have a significant effect on the profitability of the

■ 20 Banks in Germany not only own equity stakes in nonfinancial firms, but they also have proxy rights to vote the shares of other agents who keep their shares on deposit at the bank. See Mühlbert (1997) for a description of the German banking law on the monitoring instruments available to banks.

■ 21 A firm is considered to be bank-influenced if a national bank or a national insurance company holds more than 50 percent of the firm's outstanding shares or if these institutions hold more than 25 percent of outstanding shares and no other owner holds more than 25 percent.

firms with which they are associated. Wenger and Kaserer (1997) find an even less favorable result. They find that bank-dominated firms (firms in which banks have at least a 10 percent equity stake), have a lower shareholder return than non-bank-dominated firms.

Finally, some studies have looked at the impact of bank equity stakes on the performance of financially distressed firms. Hoshi, Kashyap, and Scharfstein (1990) find that the costs of financial distress are lower for Japanese firms in industrial groups than for other firms. In particular, firms in industrial groups invest and sell more than nongroup firms in the years following the onset of financial distress. They also find that firms that are not part of groups—but have close ties to a main bank—invest and sell more than firms without strong bank ties.<sup>22</sup> James (1995) identifies the conditions under which American banks take equity positions in debt restructurings. He finds, among other things, that firms in which banks take equity positions are more cash-flow-constrained and have poorer operating performance prior to the restructuring. However, these firms perform better after the restructuring than firms with no bank stock ownership.

As this sample of research shows, bank equity positions in nonfinancial firms have many effects that go beyond the potential implications on banks, though the latter tend to dominate the debate on mixing banking with commerce. Some of that research finds conflicting results, and the robustness of some of its results can be questioned. Moreover, our knowledge about many of the potential effects of banks' equity stakes in firms is still limited. For example, the empirical research conducted so far focuses on bank stakes in well-established firms. It would also seem important to ascertain the role of those stakes throughout a firm's life, in particular during the earlier stages of its life. Despite these gaps, it would seem important for policymakers to take this literature into account when debating future U.S. regulation of banking and commerce.

#### IV. Final Remarks

Throughout American banking history, regulations have generally prohibited banks from making investments in nonfinancial firms but allowed these to own equity stakes in banks. There were several occasions when banks were allowed to own stakes in firms, but those occurred mainly before the National Banking Act of 1864. The regulation of firms' ownership of banks has also become more restrictive with time, particularly since the enactment of the Bank Holding Company Act in 1956. As a result of these trends, the United States currently regulates the association between banking and commerce significantly more restrictively than other countries.

The most recent debate over loosening the current U.S. barriers that separate banking from commerce, particularly those that limit banks' and BHCs' investments in firms, has, as in the past, focused on the potential implications of this regulatory change on banks and the safety net. While undoubtedly important, these are only part of the potential effects of that association. As the literature reviewed here shows, banks' investments in firms have implications that go far beyond the banking sector. That literature has produced some conflicting results and it has not addressed some questions that are pertinent to the debate. Nonetheless, it would appear important for policymakers participating in the debate to consider that literature and, through it, the potential impact of the association between banking and commerce on the nonfinancial sector of the economy, together with the potential impact on the banking sector and the safety net.

■ 22 Sheard (1989) describes, for various Japanese firms in financial distress, the financial assistance measures adopted by the firm's main bank during restructurings.

## References

- Allen, Franklin, and Douglas Gale.** *Comparative Financial Systems: Competition Versus Insurance*, forthcoming, 1999.
- . “Financial Markets, Intermediaries, and Intertemporal Smoothing,” *Journal of Political Economy*, vol. 105, no. 3 (June 1997), pp. 523–546.
- Barth, James R., Daniel E. Nolle, and Tara N. Rice.** “Commercial Banking Structure, Regulation, and Performance: An International Comparison,” Office of the Comptroller of the Currency, Working Paper No. 97-6, 1997.
- Berlin, Mitchell, Kose John, and Anthony Saunders.** “Bank Equity Stakes in Borrowing Firms and Financial Distress,” *Review of Financial Studies*, vol. 9, no. 3 (Fall 1996), pp. 889–919.
- Bhattacharya, Sudipto, and Anjan V. Thakor.** “Contemporary Banking Theory,” *Journal of Financial Intermediation*, vol. 3, no. 1 (October 1993), pp. 2–50.
- Blair, C. E.** “Bank Powers and the Separation of Banking from Commerce: An Historical Perspective,” Federal Deposit Insurance Corporation, *Banking Review*, vol. 7, no. 1 (1994), pp. 28–38.
- Boot, Arnoud W. A., and Anjan V. Thakor.** “Financial System Architecture,” *Review of Financial Studies*, vol. 10, no. 3 (Fall 1997), pp. 693–733.
- Boyd, John H., Chun Chang, and Bruce D. Smith.** “Moral Hazard Under Commercial and Universal Banking,” *Journal of Money, Credit, and Banking*, vol. 30, no. 3 (August 1998), pp. 426–468.
- Cable, John.** “Capital Market Information and Industrial Performance: The Role of West German Banks,” *Economic Journal*, vol. 95, no. 377 (March 1985), pp. 118–132.
- Chirinko, Robert S., and Julie A. Elston.** “Banking Relationships in Germany: Empirical Results and Policy Implications,” Federal Reserve Bank of Kansas City, Working Paper 96-05, 1996.
- Cumming, Christine M., and Lawrence M. Sweet.** “Financial Structure of the G-10 Countries: How Does the United States Compare?” Federal Reserve Bank of New York, *Quarterly Review*, vol. 12, no. 4 (Winter 1987), pp. 14–25.
- Dewatripont, Mathias, and Eric Maskin.** “Credit Efficiency in Centralized and Decentralized Economies,” *Review of Economic Studies*, vol. 62, no. 4 (October 1995), pp. 541–555.
- Elston, Julie A.** “Firm Ownership Structure and Investment: Theory and Evidence from German Panel Data,” Wissenschaftszentrum Social Science Research Center, Berlin, Discussion Paper No. 93-27, 1993.
- Fein, M. L., and M. M. Faber.** Appendix to the statement of Paul A. Volcker before the Commerce, Consumer, and Monetary Affairs Subcommittee on Government Operations, U.S. House of Representatives, Washington, D.C., June 11, 1986.
- Flath, David.** “Shareholding in the Keiretsu, Japan’s Financial Groups,” *Review of Economics and Statistics*, vol. 75, no. 2 (May 1993), pp. 249–257.
- Freixas, Xavier, and Jean Charles Rochet.** *Microeconomics of Banking*. Cambridge, Mass.: MIT Press, 1997.
- Gerschenkron, Alexander.** *Economic Backwardness in Historical Perspective*. Cambridge, Mass.: Harvard University Press, 1962.
- Gorton, Gary, and Joseph G. Haubrich.** “Bank Deregulation, Credit Markets, and the Control of Capital,” *Carnegie-Rochester Conference Series on Public Policy*, vol. 26 (Spring 1987), pp. 289–334.
- and **Frank A. Schmid.** “Universal Banking and the Performance of German Firms,” National Bureau of Economic Research, Working Paper No. 5453, February 1996.
- Halpert, Stephen K.** “The Separation of Banking and Commerce Reconsidered,” *Journal of Corporation Law*, Winter 1988, pp. 481–533.

- Haubrich, Joseph G., and João A. C. Santos.** "Banking and Commerce: A Liquidity Approach," Bank for International Settlements, 1998 (unpublished).
- Hoshi, Takeo, Anil K. Kashyap, and David Scharfstein.** "Corporate Structure, Liquidity, and Investment: Evidence from Japanese Industrial Groups," *Quarterly Journal of Economics*, vol. 106, no. 1 (February 1991), pp. 33–60.
- , ———, and ———. "The Role of Banks in Reducing the Costs of Financial Distress in Japan," *Journal of Financial Economics*, vol. 27, no. 1 (September 1990), pp. 67–88.
- Huertas, Thomas F.** "Can Banking and Commerce Mix?" *Cato Journal*, vol. 7, no. 3 (Winter 1988), pp. 743–762.
- . "The Union of Banking and Commerce in American History," Appendix B to statement of Hans Angermueller, U.S. House of Representatives, December 17, 1986.
- James, Christopher.** "When Do Banks Take Equity in Debt Restructurings?" *Review of Financial Studies*, vol. 8, no. 4 (Winter 1995), pp. 1209–1234.
- John, Kose, Teresa A. John, and Anthony Saunders.** "Universal Banking and Firm Risk Taking," *Journal of Banking and Finance*, vol. 18, no. 2 (March 1994), pp. 307–323.
- Kim, Sun Bae.** "Corporate Financing Through a Shareholder Bank: Lessons from Japan," Federal Reserve Bank of San Francisco, Working Paper 92-03, 1992.
- King, B., and R. Levine.** "Financial Intermediation and Economic Development," in Colin Mayer and Xavier Vives, eds., *Capital Markets and Financial Intermediation*. Cambridge: Cambridge University Press, pp. 156–189.
- Kroszner, Randall S.** "Restructuring Bank Regulation: Conflicts of Interest, Stability, and Universal Banking," Paper presented at *Restructuring Regulation and Financial Institutions*, Milken Institute, Santa Monica, Calif., March 16, 1998.
- Langohr, Herwig, and Anthony M. Santomero.** "The Extent of Equity Investment by European Banks," *Journal of Money, Credit, and Banking*, vol. 17, no. 2 (May 1985), pp. 243–252.
- Mayer, Colin.** "New Issues in Corporate Finance," *European Economic Review*, vol. 32, no. 5 (June 1988), pp. 1167–1188.
- Mester, Loretta J.** "Banking and Commerce: A Dangerous Liaison?" Federal Reserve Bank of Philadelphia, *Business Review*, May 1992, pp. 17–29.
- Mülbert, Peter O.** "Bank Equity holdings in Nonfinancial Firms and Corporate Governance: The Case of German Universal Banks," Universität Trier, 1997 (unpublished).
- Myers, S. C., and Raghuram G. Rajan.** "The Paradox of Liquidity," *Quarterly Journal of Economics*, vol. 113, no. 3 (August 1998), pp. 733–771.
- Organisation for Economic Co-operation and Development.** *Bank Profitability: Financial Statements of Banks*. Paris: OECD, 1997.
- Pollard, Alfred M., and others.** *Banking Law in the United States*. Boston: Butterworth Legal Publishers, 1988.
- Pozdena, Randall J.** "Why Banks Need Commerce Powers," Federal Reserve Bank of San Francisco, *Economic Review*, Summer 1991, pp. 18–30.
- Prowse, Stephen D.** "Institutional Investment Patterns and Corporate Financial Behavior in the United States and Japan," *Journal of Financial Economics*, vol. 27, no. 1 (September 1990), pp. 43–66.
- Puri, Manju.** "Conflicts of Interest, Intermediation, and the Pricing of Underwritten Securities," Stanford University Graduate School of Business, Research Paper No. 1383, March 1996.
- Rajan, Raghuram G.** "The Entry of Commercial Banks into the Securities Business: A Selective Survey of Theories and Evidence," in I. Walter and A. Saunders, eds., *Universal Banking: Financial System Design Reconsidered*. Chicago: Irwin, 1996, pp. 282–302.

- . “Conflict of Interest and the Separation of Commercial and Investment Banking,” Graduate School of Business, University of Chicago, 1992 (unpublished).
- Roe, Mark J.** “Political and Legal Restraints on Ownership and Control of Public Companies,” *Journal of Financial Economics*, vol. 27, no. 1 (September 1990), pp. 7–41.
- Sabani, Laura.** “Market-Oriented Versus Bank-Oriented Financial Systems: Incomplete Contracts and Long-term Commitments,” *Journal of International and Comparative Economics*, vol. 3, (1994), pp. 279–307.
- Santos, João A. C.** “Securities Units of Banking Conglomerates: Should Their Location Be Regulated?” *Cato Journal*, vol. 18, no. 1 (Spring/Summer 1998a), pp. 93–117.
- . “Commercial Banks in the Securities Business: A Review,” *Journal of Financial Services Research*, vol. 14, (1998b), pp. 35–59.
- . “Mixing Banking with Commerce: A Review,” Bank for International Settlements, 1998c (unpublished).
- . “Bank Capital and Equity Investment Regulations,” *Journal of Banking and Finance*, forthcoming 1999.
- Saunders, Anthony.** “Banking and Commerce: An Overview of the Public Policy Issues,” *Journal of Banking and Finance*, vol. 18, no. 2 (March 1994), pp. 231–254.
- Seward, James K.** “Corporate Financial Policy and the Theory of Financial Intermediation,” *Journal of Finance*, vol. 45, no. 2 (June 1990), pp. 351–77.
- Sheard, Paul.** “The Main Bank System and Corporate Monitoring and control in Japan,” *Journal of Economic Behavior and Organization*, vol. 11, no. 3 (May 1989), pp. 399–422.
- Shull, Bernard.** “Banking and Commerce in the United States,” *Journal of Banking and Finance*, vol. 18, no. 2 (March 1994), pp. 255–270.
- . “The Separation of Banking and Commerce: Origin, Development and Implications for Antitrust,” *Antitrust Bulletin*, vol. 28, no. 1 (1983), pp. 255–279.
- Thakor, Anjan V.** “The Design of Financial Systems: An Overview,” *Journal of Banking and Finance*, vol. 20, no. 5 (June 1996), pp. 917–948.
- Weinstein, David E., and Yishay Yafeh.** “On the Costs of a Bank-Centered Financial System: Evidence from the Changing Main Bank Relations in Japan,” *Journal of Finance*, vol. 53, no. 2 (April 1998), pp. 635–672.
- Wenger, Ekkehard, and Christoph Kaserer.** “German Banks and Corporate Governance: A Critical View,” Paper presented at the *Symposium on Comparative Corporate Governance: The State of the Art and Emerging Research*, held at the Max Planck Institute for International Law, Hamburg, Germany, May 15, 1997.