

ECONOMIC REVIEW

1994 Quarter 3

**A Conference on Federal
Credit Allocation** **2**

by Joseph G. Haubrich and James B. Thomson

**Employment Creation
and Destruction:** **14**

An Analytical Review

by Randall W. Eberts and Edward B. Montgomery

**A Monte Carlo Examination
of Bias Tests in Mortgage Lending** **27**

by Paul W. Bauer and Brian A. Cromwell



FEDERAL RESERVE BANK
OF CLEVELAND



1994, Quarter 3
Vol. 30, No. 3

A Conference on Federal Credit Allocation

2

by Joseph G. Haubrich and James B. Thomson

In October 1993, the Federal Reserve Bank of Cleveland and the *Journal of Money, Credit, and Banking* sponsored a conference that examined the costs, causes, and consequences of credit allocation by the federal government. The eight presenters looked at the broad rationale for government intervention in U.S. credit markets, analyzed some issues related to pensions and federal pension guarantees, and discussed a number of specific programs and regulations, including credit imperfections in housing markets, risk-based capital requirements for banks, and community reinvestment rules. This article is an overview of those proceedings.

Employment Creation and Destruction: An Analytical Review

14

by Randall W. Eberts and Edward B. Montgomery

The capacity of markets to create jobs is typically measured by net employment changes. However, net job flows veil the dynamics underlying these aggregate figures. Recent studies have examined the cyclical behavior of the four components of net employment: jobs gained from business openings and expansions and jobs lost from business closings and contractions. This paper extends the inquiry to examine whether these employment components follow similar patterns across regions. The evidence indicates quite different behavior across regions than over time. Regional employment changes are primarily associated with job creation, whereas cyclical employment changes are associated with job destruction. Thus, policymakers need to differentiate between programs that stimulate regional job growth and those that help firms survive economic downturns.

A Monte Carlo Examination of Bias Tests in Mortgage Lending

27

by Paul W. Bauer and Brian A. Cromwell

Despite three years of data from the Home Mortgage Disclosure Act (HMDA) indicating that the rejection rate for black mortgage applicants is much higher than for whites, most financial institutions have received regulatory compliance ratings of satisfactory or better. This result may stem from the absence of several key individual characteristics in the HMDA data, which can cause tests to find bias even when it does not exist. Here, the authors examine the steps involved in determining whether a financial institution discriminates against minorities and construct a simulation model to gauge how well some of these tests perform when the degree of bias is known. They find that for plausible levels of bias, the sample size is critical, but that low levels of bias can be difficult to detect even with large sample sizes.

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 and Community Affairs Department. Call 1-800-543-3489, 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.

Coordinating Economist:
Jagadeesh Gokhale

Advisory Board:
Charles T. Carlstrom
Ian Gale
Joseph G. Haubrich

Editors: Tess Ferg
Robin Raliff
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 editors.

ISSN 0013-0281

A Conference on Federal Credit Allocation

by Joseph G. Haubrich and James B. Thomson

Joseph G. Haubrich is an economic advisor and James B. Thomson is an assistant vice president and economist at the Federal Reserve Bank of Cleveland.

Introduction

It is 10:00 a.m., and Paul Davidson is telling the assembled economists how to pull rabbits out of a hat. Metaphorically, of course: Professor Davidson was discussing the assumptions behind a thought-provoking paper presented at the Conference on Credit Allocation: Theory, Evidence, and History, held last October 17–19 in Cleveland.

The Federal Reserve Bank of Cleveland and the *Journal of Money, Credit, and Banking* sponsored this conference to support research into the costs, causes, and consequences of credit allocation by the federal government.¹ It is one of those peculiar paradoxes that federal credit allocation remains an esoteric topic despite the general familiarity with student loans, deposit insurance, and the Federal Reserve's influence on interest rates. The conference aimed to bring together an emerging body of work that looks at these issues from the standpoint of modern economics, emphasizing both common concerns and methodological differences to highlight an area that deserves greater attention.

In a world of scarce resources, *something* must allocate credit, be it the marketplace or the government. In a perfect world, prices are the most efficient method of accomplishing this. In a world with market imperfections, including significant information costs, the government might improve upon the market allocation system. However, public choice theory reminds us that government imperfections can lead to credit market intervention that reduces society's welfare. Therefore, it is important that we understand the nature of the market imperfections and the alternative solutions.

Understanding the actual effects — intended or otherwise — of particular programs is perhaps the most important immediate goal. But in thinking about the future, and in removing the prejudice about what seems “natural” or “politically feasible,” both critical theory and historical studies have a place.

The papers reviewed here were grouped into three sessions at the conference. The first presented a general overview of the problem by examining the broad rationales for credit allocation: abstract market defects and the concrete historical record. The second session took a closer look at specific programs and regulations. Housing, bank capital requirements, and

■ 1 The full proceedings appear in the August 1994 issue of the *Journal of Money, Credit, and Banking*, vol. 26, no. 3, part 2.

community reinvestment were examined analytically, empirically, and as the outcome of a political process. The third session focused on a number of issues related to pensions and federal pension guarantees. The sheer amount of money tied up in pensions makes the consequences of inept policy particularly severe — and the need for research correspondingly great.

I. Session 1: What Can Be, Might Be

Stephen Williamson first examines whether information problems justify government credit allocation. Ronnie Phillips then looks at the debate over credit allocation during the New Deal, when many current programs were first introduced and many more radical proposals were seriously considered. Finally, Marvin Goodfriend stresses the credit allocation inherent in current Federal Reserve and Treasury practices. Taken together, the papers produce a strong sense of “what can be, might be” — for better or worse.

Williamson

In “Do Informational Frictions Justify Federal Credit Programs?” Stephen Williamson evaluates the effectiveness of government credit programs using two models with imperfect information. In these models, informational frictions lead to a credit market with many realistic features, such as bonds, banks, and bad borrowers, and with flaws, such as credit rationing.

The first model looks at an economy where people must bear a cost to learn the true state of the world. (These models are hence known as “costly state verification” models.) This provides a motivation for debt and for credit rationing. Increasing the interest rate on a loan, for example, reduces the chance that the borrower will make those higher payments and thus boosts the expected verification costs. This rations some borrowers out of the market, because offering to pay a higher interest rate will not get banks to lend to them. The market treats identical borrowers differently; some get credit and some do not, even though the overall return on their investment would exceed the market interest rate.

The apparent market flaw, however, does not immediately imply that a governmental solution exists. Williamson shows that if the government credit program breaks even, neither lenders nor borrowers profit. Consider a credit guarantee

program that the government funds by charging lenders an insurance premium. With a hike in the interest rate, the bank does not bear the full increased cost of default directly, so banks overall charge a higher rate. They do bear the higher cost in the form of heftier insurance premiums, however, so their expected return can fall. Borrowers then face a steeper interest rate, while lenders get a lower expected return. This lower return means that lenders supply less capital and credit rationing gets worse. In other words, everybody loses.

The second model looks at a market where lenders must screen out bad borrowers. As lenders deny some borrowers credit, the government has a potential concern. As in the first model, however, subtle perverse effects arise from government credit allocation. Without government intervention, lower-quality borrowers, who face higher interest rates, never try to pass themselves off as high-quality borrowers because, if discovered, they get no loan at all. If the government offers loans to people who have been denied credit, it decreases the penalty for those who misrepresent their type. This in turn raises lenders’ screening costs, exacerbating the credit problem as more resources get used up in overhead and fewer are available for borrowing.

On one level, Williamson’s results may seem obvious: An unfettered market provides the best possible contracts for borrowing and lending. More important, the paper rebuts the oft-heard charge that market imperfections create a need for government intervention. Williamson goes well beyond such general issues, however, and shows that government intervention is not only unnecessary but also may prove harmful.

Paul Davidson criticizes the entire tradition behind the Williamson paper, that of classical theoretical economics. For example, he points out and questions the statistical assumptions regarding risk that allow inferences about future default rates.² Davidson also questions the model’s informational assumptions: Could private markets efficiently uncover the information and replace the implicit insurance of banks with an explicit form? Is it true that individuals know their default risk better than lenders do?

■ 2 Technically speaking, the stochastic process must be ergodic, with the time average of past values converging to the phase average across states (see Breiman [1968], chapter 6). Strictly speaking, Williamson avoids this problem because agents in this model have direct knowledge of the relevant probabilities. In actual practice, however, people must learn this from experience, which again raises the question of ergodicity.

The Davidson critique boils down to two central concerns: 1) Which assumptions best capture the real world (that is, do credit market imperfections arise from imperfect information [Williamson] or non-ergodicity [Davidson]), and 2) What vital elements has Williamson left out (competing lenders? multiple loans?). These are hard questions that are well worth thinking about, but the Williamson paper, by clearly and cleverly drawing conclusions from a well-specified set of assumptions, survives the criticism as one important way to proceed.

Pure logic cannot settle such disputes. Empirical evidence can't either, but it can help. Recent work by Berger and Udell (1992) finds little evidence of credit rationing. The classic work of Ellsberg (1961) indicates that people's perception of risk may be based more on the fear of a vague "uncertainty" than on a statistical calculation of probabilities.

Phillips

In "An End to Private Banking: Early New Deal Proposals to Alter the Role of the Federal Government in Credit Allocation," Ronnie Phillips documents that the financial reformers of the early New Deal had a list of concerns that are still voiced in policy circles today. Banks taking on riskier loans to increase profits, excessively harsh bank exams, small businesses starved for loans, Federal Reserve accountability, and the incentive effects of deposit insurance were only some of the topics the reformers considered. Not surprisingly, many policy prescriptions of that era also look familiar, encompassing narrow banks, a Reconstruction Finance Corporation, and changes in the Fed's discount window policy.

Phillips describes how the politicians and economists of the time clearly understood the important distinction between money and credit. Credit is the transfer of real resources from lenders to borrowers; money is the means of payment, or the medium of exchange. The proposals seriously considered 100 percent reserve banking, establishing small mutual savings associations for local lending, and extensions of the Reconstruction Finance Corporation to allocate credit globally among local associations and between larger corporations. One set of lessons from the New Deal debates that deserves more emphasis is the objections people had to the proposals — the fear of political control, evidence of ineffectiveness, and lack of specific objectives.

A less direct lesson concerns the importance of history and the path-dependent nature of economic experience. Phillips strongly suggests that a series of historical accidents lay behind the New Deal's rejection of the proposals to separate money and credit. Roosevelt's failure to consult Senator Carter Glass before appointing Marriner Eccles to the Federal Reserve chairmanship led to Glass' obstructing key reforms. A prominent senator supporting the reforms, Bronson Cutting of Nevada, died in a plane crash before the measures could be put to a vote. These facts strongly suggest that our current system has arbitrary components and deserves a fresh look.

Walker Todd's discussion heightens the relevance of the historical perspective provided by Phillips. He brings out the larger political economy issues framing the 1930s' debate, showing how the early New Deal reforms were opposed both by those advocating increased government involvement and by traditional fiscal conservatives. The continuing importance of these strands of thought means, paradoxically, that the historical record can provide a useful commentary on current proposals.

Goodfriend

Marvin Goodfriend provides a natural follow-up to the history lesson, arguing persuasively that we live with the adverse consequences of not facing up to the distinction between money and credit. In "Why We Need an 'Accord' for Federal Reserve Credit Policy," he suggests extending the 1951 Fed-Treasury accord, which eliminated the central bank's commitment to support government bond prices. Goodfriend similarly wants to free current Fed credit operations from potential abuse because of concerns over fiscal policy.

The author defines monetary policy as a change in the stock of high-powered money, while credit policy is a change in the central bank's assets that keeps the stock of high-powered money fixed. In his view, an effective central bank should not be distracted by entanglements peripheral to its mission. The accord of 1951 effectively freed Federal Reserve monetary policy from entanglement with fiscal problems, but the necessary credit actions of the central bank (especially as the lender of last resort) currently have little or no protection. Applying three basic principles would provide such protection: Limit assistance to illiquid but solvent institutions, do not

use credit policy to avoid congressional authorization of funding, and do not directly transfer Fed surplus to the Treasury.

Applying these rules would have some non-trivial implications. Quite clearly, it would prevent discount window lending from delaying the closure of insolvent banks (see Todd [1988, 1993]). Less obviously, it would also affect foreign exchange intervention. When the Federal Reserve buys German marks, for example, it acquires international reserves that increase the money supply; this constitutes monetary policy. But usually, the Fed acts to “sterilize” the intervention, offsetting the increase in reserves by selling domestic securities: Sterilized intervention thus constitutes credit policy. Goodfriend wants this type of intervention to be subject to explicit congressional authorization. Similarly, he believes that Congress should authorize “warehousing,” wherein the Fed buys foreign currency from the Treasury and agrees to sell it back at some point in the future. In both cases, managing the balance-of-payment deficits constitutes proper fiscal policy actions of the U.S. government. Central bank activity obscures the funding process and would fall under the proposed accord.

The “credit accord” would also stop the transfer of Federal Reserve surplus to the Treasury (authorized by the Deficit Reduction Act of 1993). The Fed surplus is part of its capital account and represents retained earnings. Traditionally, the central bank has maintained a surplus account equal to paid-in capital. Goodfriend explains how the transfer, which the Fed finances by selling Treasury securities in its portfolio, results in no actual deficit reduction in the long run. That’s because once the securities are sold to the public, the Fed no longer remits to the Treasury any interest earned on Treasury securities. Over time, the loss of this revenue offsets the surplus transferred.

E.J. Stevens’ discussion traces the problems that Goodfriend seeks to resolve to an even deeper source: the lack of clear objectives for Federal Reserve policy. Given a pessimism about any near-term change in this situation, an accord may serve as a second-best way to extricate the central bank from inappropriate transactions. This immediately raises two questions: Which transactions are inappropriate, and who should be party to the accord?

Stevens argues that in most cases, the Fed is not the most appropriate party to the accord. Under the Federal Deposit Insurance Corporation Improvement Act of 1991 (FDICIA), an accord already exists limiting Fed lending to insolvent

institutions; thus, a tune-up may be desired, but is a major reform in order? Likewise, the Treasury and Congress can *by themselves* refrain from using the Fed for underhanded financing, be it by foreign exchange warehousing or the transfer of surplus funds.

Ultimately, the issue is that while Congress and the Treasury may not want a commitment to avoid entangling the Fed, the Fed may desire such a commitment. It really does dieters no good to know that if all ice cream manufacturers voluntarily ceased production, they could stay thin. The more difficult issue that Goodfriend and Stevens wrestle with is the actual importance of the particular problems Goodfriend cites.

II. Session 2: Specific Programs

The second session of the conference looks at specific government programs in more detail. Allen Berger and Greg Udell evaluate risk-based capital requirements for banks. Since such requirements alter the relative cost of funding for different types of assets, the policy may have credit allocation consequences. Charles Calomiris, Charles Kahn, and Stanley Longhofer then model the credit imperfections in housing markets and assess the possibility of beneficial government intervention. Finally, Anjan Thakor and Jess Beltz look at the political economy behind rules that target bank credit toward specific groups.

Berger and Udell

In “Did Risk-Based Capital Allocate Bank Credit and Cause a ‘Credit Crunch’ in the United States?” Allen Berger and Gregory Udell investigate the reallocation of bank credit from loans to securities in the early 1990s. In searching for the cause of this portfolio shift, they test a variety of possibilities, including the imposition of risk-based capital requirements, tougher loan examinations, increased leverage requirements, and several nonregulation-based reasons. Berger and Udell are the first to examine all of these competing theories simultaneously.

The study utilizes an extensive data set covering quarterly numbers on almost all U.S. commercial banks between 1979 and 1992. It is also distinguished by the use of a control period that lets the authors determine if the early 1990s look different enough to merit designation as a “credit crunch.” To do this, Berger and Udell estimate a series of supply equations

for bank credit and then test for differences in credit behavior between the crunch period and the control period.

They also compare the portfolio allocation decisions of well-capitalized banks with those of undercapitalized banks, since a number of the hypotheses predict differences in portfolio allocation effects across these subsamples.

How the supply equations shift allows Berger and Udell to distinguish between the competing theories.

The results provide little support for most of the supply-side theories of bank portfolio shifts in the early 1990s. For example, the estimated supply equations show that banks with low risk-based capital ratios did not reduce their lending or increase their securities between the control period and credit crunch more than did banks with higher risk-based capital ratios. The other supply-side stories provide a better explanation than risk-based capital, but still seem quantitatively unimportant. The demand-side theories fare better, though perhaps only because their effects are harder to pin down with banking data.

Berger and Udell's conclusions differ from previous work showing that risk-based capital requirements significantly alter banks' portfolio behavior, such as Haubrich and Wachtel (1993). Much of the discrepancy probably can be traced to different empirical methods. Berger and Udell stress the importance of the control period and of assessing a credit crunch only in relation to the control. By claiming that a credit crunch occurs *only* if the loan supply function differs between the crunch and the control, the paper drives this point too far. The imposition of risk-based capital requirements has increased the number of capital-constrained banks; these institutions then reacted to the constraints. The behavior of a capital-constrained bank during the control period, however, need not differ from the actions of one facing risk-based standards. Tough regulators could have caused the problem either because they got tougher or because banking conditions exposed their inherent toughness. This is really the old economic distinction between movements *along* a supply curve versus movements *of* a supply curve. A decrease in demand can cause a reduction in equilibrium quantity even if the supply curve does not shift.

Merwan Engineer provides a thoughtful commentary on the Berger and Udell paper. He correctly points out that since risk-based capital guidelines were imposed internationally, a comparison with other countries may help to

resolve the issue in dispute. He also points out that a common problem is estimating supply equations for heavily regulated industries such as banking: Heavy regulation generally means that supply relations change frequently, making it difficult to get a fix on them. In the case at hand, variables measuring bank risk influence bank portfolio choice via regulatory behavior, so when regulations change, the estimated relation should change. Finally, he argues that the issues considered fit into a broader context that was not raised in the paper: Were the new capital standards wise, and if so, were they adopted at the right time?

Calomiris, Kahn, and Longhofer

In "Housing-Finance Intervention and Private Incentives: Helping Minorities and the Poor," Charles Calomiris, Charles Kahn, and Stanley Longhofer look at the goals behind government housing programs, the possible market imperfections that may justify such intervention, and the costs and benefits of the intervention. Economists typically fall back onto equity and efficiency issues when undertaking such an analysis of government programs. Unfortunately, while equity is the main motivator of government housing intervention, equity (unlike economic efficiency) is a slippery concept with different meanings for different individuals. Consequently, even though efficiency is the main yardstick for examining federal housing intervention, Calomiris et al. provide the reader with a framework for understanding the equity issues. They define three types of equity — equitable procedures, equal outcomes, and retributive justice — and illustrate how each may imply a different form of intervention or program design.

After providing an overview of the different types of housing interventions, the authors outline four different classes of housing market problems that government interventions could be designed to solve: wealth inequality and poverty, informational externalities, bigotry, and rational discrimination. Although the issues and relevant literature on each class of problem are reviewed, the presentation deals mainly with rational discrimination.

Rational discrimination in housing finance arises from informational asymmetries associated with evaluating mortgage applications. At each stage of the credit evaluation of an applicant, the lender must determine whether to

continue collecting information or to stop (that is, to deny the loan). Calomiris et al. argue that if nonminority loan officers find it harder to interpret signals from minorities, then they must make the loan relying on fewer informative signals, and the bank both denies a larger share of minority loan applications and faces a higher degree of minority default. As a result, the lender's costs of processing minority applications are higher and its expected return is lower. Therefore, lenders will employ more conservative stopping rules to minority applications.

This is obviously discrimination, since otherwise identical minority and nonminority applicants face different probabilities of receiving credit. It is rational because the higher denial rate is not a function of bigotry, but rather is based on the lender's private benefits and costs associated with information collection.

The rational discrimination model is particularly interesting because it resolves the "Becker paradox." If minorities face discrimination, the paradox runs, they should have lower default rates than whites, since banks lend only to ultra-safe minority borrowers. This contradicts the evidence, which shows that minorities have higher default rates than do whites, even after controlling for all of the relevant economic variables. The empirical evidence, however, is consistent with Calomiris et al.'s rational discrimination model.

In a review of some recent evidence on discrimination in mortgage markets, the authors find that most of the formal and informal evidence is consistent with rational discrimination. Moreover, they argue that explanations such as cultural affinity or bigotry cannot explain the poor Community Reinvestment Act (CRA) ratings of minority-owned banks. However, to the extent that minorities face educational disadvantages and have lower average wealth than society in general, the poor CRA ratings of minority-owned firms are not inconsistent with rational discrimination.

If what we observe in housing finance markets is indeed rational discrimination, then what is the appropriate policy response by government? Calomiris et al. conclude that subsidized community development banks appear to be the most efficient solution.

Robert Van Order's critique of the Calomiris paper provides some useful insights on both the issue of discrimination and the analysis of discrimination. Van Order questions the usefulness of empirical studies of mortgage discrimination because of the omitted-variables problem. If explanatory variables omitted from the regression experiment are correlated with race, biases are introduced in the race coefficient of the logit

regression. Therefore, considerably more work is needed before strong conclusions can be drawn from these studies. Van Order also questions conclusions about the "Becker paradox" drawn from mortgage default rates. He points out that the proposition is based on the default rate of the marginal borrower, not the average one, and that the "econometric problems of isolating what is marginal are formidable."

Van Order suggests that the model of rational discrimination is consistent with commission-based compensation for lending officers. When commissions are based on the number of applications processed that meet underwriting standards, lending officers find efficient ways to allocate their time. These include stopping rules based on variables that have been found to be correlated with the creditworthiness of the borrower, including race. Van Order argues that rational discrimination can be dealt with through the same testing and enforcement mechanisms used to counteract bigotry.

He also makes a useful distinction between community lending issues and issues of discrimination. Distinguishing between these two is important because the solutions may be quite different. On one hand, community lending issues are about channeling funds into depressed areas. Van Order notes that the problem here is not necessarily one of race but rather of neighborhood externalities. On the other hand, he suggests that discrimination is most serious for middle-class blacks, most of whom do not live in depressed areas. In his view, one of the problems with CRA is that it does not distinguish between these two sometimes conflicting issues and therefore does a poor job of solving either one.

Thakor and Beltz

In "A 'Barter' Theory of Bank Regulation and Credit Allocation," Anjan Thakor and Jess Beltz advance the discussion of government involvement in credit markets beyond market failure and instead attempt to understand the self-interested motives behind these interventions. They posit that the existing complex web of regulatory subsidies and taxes is the outcome of what starts out as a mutually beneficial barter arrangement.

In their model, government subsidies benefit banks more than they cost taxpayers. In return, banks allow the government to dictate some aspects of their credit allocation. For example, they may accept the CRA in exchange

for deposit insurance and access to the discount window. This barter arrangement is a dynamic one, however, with the costs of government intervention increasing as new regulations are needed to counteract banks' circumvention of the original statutes.

Such a regulatory exchange can be mutually beneficial as long as the subsidy exceeds the cost of regulation. If a subset of banks finds this barter arrangement unprofitable, then a bad Nash equilibrium arises, which Thakor and Beltz call a regulatory trap. A bank opting out of the system would not attract any depositors, who prefer institutions covered by deposit insurance. This holds true as long as some banks find the regulatory barter process to be profitable. The banking system as a whole may prefer to give up deposit insurance when credit allocation regulations become too imposing, but a coordination problem prevents this.

Thakor and Beltz's basic model is a two-period one in which all agents are risk averse. There are three types of borrowers: the good (G), the bad (B), and the underprivileged (U). Only G borrowers have positive net-present-value projects in which to invest. The bank cannot distinguish between G and B borrowers, but it can readily identify the U's. Without government intervention, banks would never lend to U borrowers.

All projects last one period, and banks can infer the borrower type by the realized return on the first-period project. Only G borrowers with successful first-period projects obtain financing in the second period. Finally, successful G borrowers in the second period may have only a risky project to invest in or a choice between a risky project and a higher-valued safe one. However, only the lending bank in the first period can distinguish between borrowers with and without a choice of second-period projects. This is an important assumption because depositors will price their deposits according to the perceived attributes of the bank's portfolio. The inability of depositors to see the banker's private information and the incentives banks have to misrepresent their information on borrowers lead to a higher deposit rate than would be obtained in the full information case. Unfortunately, at this higher deposit rate, the banker cannot offer borrowers with a project choice a lending rate that will make the safe project profitable. Hence, the agency conflict both prevents the bank from exploiting its proprietary information and distorts investment.

In this model, government provision of a financial safety net through deposit insurance is Pareto improving. Deposit insurance solves the

agency conflict between depositors and the banks arising from informational asymmetries by making deposits riskless. As a result, deposit insurance allows a bank to earn rents on its monopoly information in the second period by allowing it to price its loans according to borrower characteristics. Lower interest rates can be offered to borrowers along with an option to invest in the safe project, thus making the choice of the more highly valued safe project optimal. Consequently, deposit insurance removes the second-period investment distortion arising from informational asymmetries. Finally, if deposit insurance is underpriced, then banks and borrowers share in the surplus that results. However, given that the deposit insurance subsidy is available to banks that invest their deposits in marketable securities, a barter agreement between banks and the government becomes feasible.

The final element in Thakor and Beltz's analysis is the introduction of lending to U borrowers as a political good. Then, as a condition for insurance, the government mandates that banks lend a fixed portion of their deposits to U borrowers. Banks are willing to enter into this contract as long as the increase in profits from access to deposit insurance exceeds the cost of complying with the lending regulation.

Thakor and Beltz then show that if this barter arrangement becomes unprofitable over time for a subset of banks, these institutions will be trapped into maintaining the arrangement as long as other banks find it profitable to continue. This regulatory trap arises because the profits of the trapped banks are conditional on the actions of the nontrapped institutions. Unilaterally dropping deposit insurance would lower a bank's profits if other banks do not follow and instead retain their insurance. This occurs because the bank dropping its insurance will have higher funding costs than the insured institution and therefore will be unable to compete in the lending market. However, all banks could profit if the barter arrangement were dropped (that is, if they all canceled their insurance) and side payments were made from the trapped banks to the untrapped ones. This solution is precluded by coordination problems.

The essence of the Thakor/Beltz analysis is that social regulation and financial safety net subsidies go hand in hand. It is therefore unrealistic to argue for a reduction in the regulatory burden without a reduction in the subsidies. It is also impractical to think that one can extend social regulation to nonbank financial firms without also providing them access to deposit insurance and the discount window.

In her comments on Thakor and Beltz's paper, Deborah Lucas raises a number of valid concerns about the analysis. First, she correctly notes that the government could unilaterally impose regulations on the banking industry without offering special subsidies. Therefore, the authors need to explain why banks are different from industries such as automobile manufacturers, who face costly regulations (fuel efficiency standards, for instance) but do not appear to be compensated. In the absence of such an explanation, the barter theory seems less justified than a straight regulatory tax story.

Lucas also raises questions about the robustness of the Thakor/Beltz results to different modeling assumptions. As she notes, by introducing subsidized deposit insurance as a means of solving the monitoring problem, the authors produce an outcome in which deposit insurance lowers bank risk. This, of course, is at odds with the option-pricing approach to valuing deposit insurance and the attendant moral hazard problem, which leads to increased bank risk (Merton [1977]).

Finally, Lucas agrees with the authors' conclusions regarding the implications of the bad Nash equilibrium and that banks as a whole could benefit by opting out of the unprofitable barter arrangement. However, she points out that if the story is not one of barter but rather a simpler one of regulatory taxes, then the same policy conclusions may not apply.

III. Session 3: Pensions

The third and final session takes a more in-depth look at a particular area — pensions — where federal programs may have a huge impact on both individual fortunes and economywide variables. Indeed, the parallels between the Pension Benefit Guaranty Corporation (PBGC) and the Federal Savings and Loan Insurance Corporation (FSLIC) are at times uncanny. Like the savings and loan debacle, the large and growing contingent liabilities of the PBGC have the potential to strike a public nerve — as does the cost of any bailout. The two papers in this session offer a somewhat different perspective on an aspect of federal credit allocation: that of ascertaining the facts. The answers are not always easy to obtain, even to straightforward questions such as “How much do people contribute to their pensions?” or “How valuable is PBGC insurance?”

Gale

William Gale offers a new look at the determinants of pension contributions in “Public Policies and Private Pension Contributions,” which provides the reader with a synthesis of the literature on pensions and pension contributions. Many aspects of the contribution decision — tax deductibility, benefit guarantees, and vesting rules — depend on government regulations, especially the Employee Retirement Income Security Act of 1974 (ERISA). Uncovering the smoking gun that links shifts in policy with shifts in contributions has been difficult, in part because the data on private pension contributions are sparse. Moreover, shifts in demographics during the 1970s and 1980s further cloud the issue.

Gale provides the reader with some insights into the issues surrounding pensions by examining changes in the legal and regulatory environment as well as trends in pension coverage, pension plan choice (defined benefit versus defined contribution), and funding status over time. He then sorts out what the literature has to say about these recent trends and the role that government intervention into the nation's pension markets has played. For example, ERISA was a major force in the shift by employers from defined benefit to defined contribution plans. However, the pension literature points to changes in industrial composition and employment as two other important factors explaining the shift in plan choice.

In seeking to understand the determinants of private pension contributions, Gale is faced with isolating the effects of government intervention into the nation's pension markets using data that are fragmented and inconsistent over time. To deal with this issue, he estimates his empirical model using two different sets of data. First, he uses standard data from the National Income and Product Accounts (NIPA). Unfortunately, the NIPA data are inadequate because they omit employee contributions, which are a growing and important share of total contributions. To control for this deficiency, Gale constructs a second measure of pension contributions by piecing together the standard NIPA figures both with IRS Form 5500 reports from private pension plans and with a new Brookings Institution series. With this improved, comprehensive series in hand, real contributions per worker are correlated with earnings, per capita asset holdings, the previous year's contributions, and dummy variables for various regulatory episodes.

Gale's empirical results highlight the inadequacy of standard data sets such as the NIPA. Time-series regressions using the dependent variable constructed from NIPA data reveal no evidence that ERISA or the Omnibus Budget Reconciliation Act of 1987 (OBRA87) had any effect on real pension contributions. Yet, when the dependent variable is taken from Gale's "constructed measure," both are shown to be significant factors affecting private pension contributions. The study finds that ERISA increased annual real contributions by an average of \$213 per person and that OBRA87 reduced contributions by \$154.

One caveat on the regressions using the constructed measure of real contributions per worker is that both ERISA and OBRA87 occur close to the sample dates at which data from different sources were spliced together. Therefore, the ERISA and OBRA87 dummy variables could be proxying for rules that Gale used to arrive at his constructed measure.

In his review of Gale's work, Joseph Ritter notes that private pensions are part of the compensation packages offered by some firms and as such may be an important part of the structure of incentives used to motivate workers. In other words, there may be much more to pensions than their impact on the structure of compensation and the composition of private savings. Consequently, government policies affecting private pensions may have important spillover effects on labor and capital markets.

Ritter ultimately finds the paper to be interesting, well motivated, and a useful survey of data sources, legal and regulatory changes, and empirical evidence. Moreover, he finds the empirical tests to be appropriate and well executed. He does question the robustness of the results, however, because the dependent variable (employee composition) is constructed four different ways across time. Unfortunately, the shifts in how employee composition is constructed tend to coincide with the events Gale is studying.

Pennacchi and Lewis

George Pennacchi and Christopher Lewis seek to determine "The Value of Pension Benefit Guaranty Corporation Insurance" by modeling PBGC guarantees as a put option with a stochastic exercise date. At first sight, this may seem like a lot of machinery for one number, but the number lies at the heart of the PBGC problem. Evaluating the PBGC's assets and liabilities lets us

know if the insurance fund is healthy, tottering, or another FSLIC waiting to explode.

Merton (1977) shows that financial guarantees like PBGC insurance can be modeled as a put option (that is, the right but not the obligation to sell a stock at a predetermined price). To value a standard put option, however, one must know the exercise price and the exercise date, which for pension guarantees are unknown. Unfortunately, while reliable estimates of the exercise price can be obtained, the exercise date cannot be predicted accurately. To resolve this problem, Marcus (1987) values PBGC liabilities as a futures contract with a maturity equal to the time of the sponsoring firm's (pension plan's) bankruptcy. This futures contract model for valuing PBGC guarantees links the value of the guarantee to both the financial condition of the pension fund and the likelihood that the sponsoring firm will become bankrupt.

Pennacchi and Lewis take a different tack. Using a continuous-time options-valuation approach, they value PBGC guarantees as a put option with an uncertain exercise date. They thus extend Marcus' model in an important way. Futures contracts are different from options because they represent an *obligation* to buy or sell an underlying asset at a future date, as opposed to the right to carry out the transaction. Consequently, Marcus' formula for valuing pension guarantees implicitly assumes that the PBGC would experience a gain whenever a bankrupt firm's pension plan was overfunded. Pennacchi and Lewis' put-option formula explicitly recognizes that the PBGC's guarantee is contingent on both a firm's bankruptcy and its pension plan being underfunded, or insolvent.

Pennacchi and Lewis add another important wrinkle to valuing PBGC guarantees. They attempt to control for the firm's ability to increase its pension liabilities in the period just preceding bankruptcy.³ To do this, they gross up the firm's pension liabilities by a factor λ .

The study shows that the value, at time zero, of the PBGC guarantee on one dollar of accrued pension liability is a positive function of λ and of the ratio of grossed-up pension liabilities to pension assets. The value of the PBGC guarantee is a negative function of the time remaining until the firm goes bankrupt.

■ 3 A firm's ability to adjust its balance sheet dynamically in response to external events is what Ritchken et al. (1993) call the *flexibility option*.

After solving for the value of a put option with a random exercise date contingent on firm bankruptcy, Pennacchi and Lewis use some representative parameter values to calculate the PBGC's liabilities and to conduct some comparative statics exercises. The exercises show that the put option model always yields higher pension costs than does Marcus' futures contract model.

For firms with low pension funding ratios and low net worth, Marcus' model appears to be a good approximation of the put. This is simply because a put option that is "deep in the money" (one that is almost certain to be exercised) is very similar to a futures contract. The bias in the Marcus model increases along with a firm's net worth. Interestingly, this is because a high level of net worth gives a firm with an overfunded pension plan time to underfund it. For similar reasons, the bias also rises with the pension-funding-to-liability ratio.

Andrew Chen's follow-up discussion points to the contingent put option model of PBGC insurance as an important contribution to the pension literature. He notes that the comparative statics performed by Pennacchi and Lewis provide useful insights into the properties of the PBGC and produce results consistent with economic intuition.

Overall, Chen finds the Pennacchi/Lewis paper to be an important contribution to the literature, but suggests that the analysis is incomplete. While Pennacchi and Lewis' model is a clear improvement of Marcus' futures contract model, Chen raises five questions about the model and its assumptions. His strongest criticisms are that Pennacchi and Lewis ignore taxes in their analysis and do not look at the volatility of pension assets. Chen argues that a complete analysis of PBGC guarantees must account for the tax factor, which is a major determinant of corporate pension asset and funding decisions. Furthermore, he suggests that the comparative statics for the volatility of pension assets must be explored.

Chen also offers some other less serious criticisms of the analysis. First, he finds the authors' use of market-value insolvency as a proxy for bankruptcy to be inconsistent with the legal definition of bankruptcy (a firm's inability to meet its contractual payments obligation). Second, he argues that the assumption underlying scaling up pension liabilities at termination by a factor λ is inconsistent with the empirical evidence found in Bodie et al. (1987), which suggests that eleventh-hour increases in pension liabilities are uncommon. Finally, he questions whether modeling PBGC guarantees as an infinite-maturity option is

superior to a one-period option with an uncertain exercise price.

IV. Conclusion

Each paper presented at the conference illuminates some important aspect of federal credit allocation. Taken together, they illustrate the range and significance of the government's intervention into the broad credit market. Some of the work has practical applications already, such as valuing and quantifying the effects of regulations. Collectively, the papers might best be thought of as a series of warnings: Some simple insights and obvious stories turn out to be untrue; some easy solutions don't work. Taken as a body, the papers also point to three unresolved issues that demand attention:

- 1) How important are racism, credit rationing, and other imperfections in the credit market?
- 2) Given credit market imperfections, what is the best means of resolving the problems — regulations, taxes and subsidies, or government organizations?
- 3) How will actual, as opposed to ideal, solutions work in the real-world political and economic environment?

The overriding normative question about the desired extent of government intervention remains open. We believe that the papers presented here provide both a direction and a springboard for needed future research.

References

- Berger, Allen N., and Gregory F. Udell.** "Some Evidence on the Empirical Significance of Credit Rationing," *Journal of Political Economy*, vol. 100, no. 5 (October 1992), pp. 1047–77.
- _____, and _____. "Did Risk-Based Capital Allocate Bank Credit and Cause a 'Credit Crunch' in the United States?" *Journal of Money, Credit, and Banking*, vol. 26, no. 3, part 2 (August 1994), pp. 585–628.
- Bodie, Zvi, et al.** "Funding and Asset Allocation in Corporate Pension Plans: An Empirical Investigation," in Z. Bodie, J. Shoven, and D. Wise, eds., *Issues in Pension Economics*. Chicago: University of Chicago Press, 1987, pp. 15–44.
- Breiman, Leo.** *Probability*. Menlo Park, Calif.: Addison-Wesley, 1968.
- 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, part 2 (August 1994), pp. 634–74.
- Chen, Andrew H.** Comment on "The Value of Pension Benefit Guaranty Corporation Insurance," *Journal of Money, Credit, and Banking*, vol. 26, no. 3, part 2 (August 1994), pp. 754–56.
- Davidson, Paul.** Comment on "Do Informational Frictions Justify Federal Credit Programs?" *Journal of Money, Credit, and Banking*, vol. 26, no. 3, part 2 (August 1994), pp. 545–51.
- Ellsberg, Daniel.** "Risk, Ambiguity, and the Savage Axioms," *Quarterly Journal of Economics*, vol. 75, no. 4 (November 1961), pp. 6453–669.
- Engineer, Merwan.** Comment on "Did Risk-Based Capital Allocate Bank Credit and Cause a 'Credit Crunch' in the United States?" *Journal of Money, Credit, and Banking*, vol. 26, no. 3, part 2 (August 1994), pp. 629–33.
- Gale, William G.** "Public Policies and Private Pension Contributions," *Journal of Money, Credit, and Banking*, vol. 26, no. 3, part 2 (August 1994), pp. 710–32.
- Goodfriend, Marvin.** "Why We Need an 'Accord' for Federal Reserve Credit Policy: A Note," *Journal of Money, Credit, and Banking*, vol. 26, no. 3, part 2 (August 1994), pp. 572–80.
- Haubrich, Joseph G., and Paul Wachtel.** "Capital Requirements and Shifts in Commercial Bank Portfolios," Federal Reserve Bank of Cleveland, *Economic Review*, vol. 29, no. 3 (1993 Quarter 3), pp. 2–15.
- Lucas, Deborah.** Comment on "A 'Barter' Theory of Bank Regulation and Credit Allocation," *Journal of Money, Credit, and Banking*, vol. 26, no. 3, part 2 (August 1994), pp. 706–09.
- Marcus, Allen.** "Corporate Pension Policy and the Value of PBGC Insurance," in Z. Bodie, J. Shoven, and D. Wise, eds., *Issues in Pension Economics*. Chicago: University of Chicago Press, 1987, pp. 49–76.
- Merton, Robert C.** "An Analytic Derivation of the Cost of Deposit Insurance and Loan Guarantees: An Application of Modern Option Pricing Theory," *Journal of Banking and Finance*, vol. 1, no. 1 (June 1977), pp. 3–11.
- Pennacchi, George G., and Christopher M. Lewis.** "The Value of Pension Benefit Guaranty Corporation Insurance," *Journal of Money, Credit, and Banking*, vol. 26, no. 3, part 2 (August 1994), pp. 735–53.
- Phillips, Ronnie J.** "An End to Private Banking: Early New Deal Proposals to Alter the Role of the Federal Government in Credit Allocation," *Journal of Money, Credit, and Banking*, vol. 26, no. 3, part 2 (August 1994), pp. 552–68.
- Ritchken, Peter, et al.** "On Flexibility, Capital Structure, and Investment Decisions for the Insured Bank," *Journal of Banking and Finance*, vol. 17, no. 6 (December 1993), pp. 1133–46.
- Ritter, Joseph A.** Comment on "Public Policies and Private Pension Contributions," *Journal of Money, Credit, and Banking*, vol. 26, no. 3, part 2 (August 1994), pp. 733–34.

Stevens, E.J. Comment on "Why We Need an 'Accord' for Federal Reserve Credit Policy: A Note," *Journal of Money, Credit, and Banking*, vol. 26, no. 3, part 2 (August 1994), pp. 581–84.

Thakor, Anjan V., and Jess Beltz. "A 'Barter' Theory of Bank Regulation and Credit Allocation," *Journal of Money, Credit, and Banking*, vol. 26, no. 3, part 2 (August 1994), pp. 679–705.

Todd, Walker F. "Lessons of the Past and Prospects for the Future in Lender of Last Resort Theory," *Proceedings of a Conference on Bank Structure and Competition*, Federal Reserve Bank of Chicago, May 11-13, 1988, pp. 533–77.

_____. "New Discount Window Policy Is Important Element of FDICIA," *Banking Policy Report*, vol. 12, no. 5 (March 1, 1993), pp. 1, 11–17.

_____. Comment on "An End to Private Banking: Early New Deal Proposals to Alter the Role of the Federal Government in Credit Allocation," *Journal of Money, Credit, and Banking*, vol. 26, no. 3, part 2 (August 1994), pp. 569–71.

Van Order, Robert. Comment on "Housing-Finance Intervention and Private Incentives: Helping Minorities and the Poor," *Journal of Money, Credit, and Banking*, vol. 26, no. 3, part 2 (August 1994), pp. 675–78.

Williamson, Stephen D. "Do Informational Frictions Justify Federal Credit Programs?" *Journal of Money, Credit, and Banking*, vol. 26, no. 3, part 2 (August 1994), pp. 523–44.

Employment Creation and Destruction: An Analytical Review

by Randall W. Eberts and Edward B. Montgomery

Randall W. Eberts is the executive director of the W. E. Upjohn Institute for Employment Research, Kalamazoo, Michigan, and Edward B. Montgomery is a professor of economics at the University of Maryland and a research associate of the National Bureau of Economic Research. The authors would like to thank Bennett Harrison and Jagadeesh Gokhale for helpful comments, and Stephen Davis and John Haltiwanger for useful discussions and for sharing their data.

Introduction

Most assessments of labor market performance over a business cycle or across regions focus on changes in net employment rates. Hidden behind the veil of these aggregate numbers are four components of employment change: jobs gained from business openings, jobs gained from business expansions, jobs lost from business contractions, and jobs lost from business closings. In the last several years, a number of studies have identified and examined these components over time and across regions to gain additional insights into the performance and dynamics of labor markets.

Labor market dynamics are characterized by two types of turnovers. One is the transition of workers into and out of positions; the second is the change in the number of jobs. While these decisions are interrelated, they are aligned with supply and demand responses. Workers move between jobs to better match their skills, wage expectations, and workplace preferences with the attributes of the position. Businesses change the number and type of employment positions in response to shifts in product demand and factor costs. Traditionally, research on labor market dynamics has concentrated on

the supply-side responses to labor market shocks by examining worker decisions to move into and out of the labor force or between employment and unemployment. This paper focuses on jobs by tracking employment changes resulting from the opening, expansion, contraction, and closing of individual establishments. Examining the components of job creation and destruction provides insight into the employment turnover process beyond what can be learned by looking only at the flow of workers.

Our objectives are twofold. First, we review previous studies of job creation and destruction to see what consensus has emerged about the demand-related side of labor market dynamics. Second, we present new evidence from several sources to augment existing evidence on differences in the causes of high- and low-frequency movements in employment. In particular, we look at whether the components of cyclical and secular (regional) variations in job growth follow similar patterns. Are fluctuations in employment over business cycles correlated more with variations in job creation (openings and expansions) than with variations in job destruction (contractions and closings)? Is employment growth in some regions characterized by greater job creation or fewer job losses?

While these two questions appear to be similar, none of the earlier studies has directly compared the behavior of cyclical and regional employment components. Our evidence suggests that these components behave quite differently over time and across regions. We find that employment fluctuations over business cycles are associated primarily with job destruction, whereas employment differences across regions are associated more with job creation.

These insights may have important policy implications at both the local and national levels. For instance, since regional employment differences are correlated more with job creation than with job destruction, state and local policies aimed at promoting new firm creation and expansion might be more fruitful in the long run than those directed toward aiding ailing firms. On the other hand, since cyclical employment is associated more with job destruction, it may be prudent to design policies to help firms through economic downturns so that fewer workers are laid off and less hardship is incurred. Clearly, definitive policy recommendations must await a more structural analysis of the determinants of job creation and destruction. Nonetheless, the results presented here may be of value in guiding this structural modeling and may serve as a cautionary note to policymakers that existing actions could be working against the economic forces that generate employment growth.

I. Definitions and Data

Studies of the demand-side components of employment change depend on longitudinal establishment-level data. By definition, an establishment is considered an opening if it did not exist at the beginning of the period but did exist at the end. A closing is defined conversely. Therefore, employment gains from openings are the sum of employment in establishments that were not present at the beginning of each period but that did exist at the end. Employment losses from closings refer to employment at those establishments that were in the data set at the beginning of the period but absent at the end. Employment shifts due to expanding or contracting firms are based on job changes at those entities that are present at both the beginning and the end of each period.

Two issues arise in constructing the data sets that could affect the relative contributions of the four components of net employment change. The first is the frequency of observations. The proportion of jobs created from openings or expansions (or lost as a result of closings or contractions) is sensitive to the length of time between the beginning and the end of the period used to construct each component. Given a time-invariant stochastic process of openings and closings, a greater proportion of employment gains would be attributed to openings than to expansions as the period between observations lengthens. To illustrate, consider the extreme case in which the time period chosen is from 1789 to the present. Here, virtually all U.S. employment would have been generated from openings. Obviously, job creation — openings and expansions combined — would not be affected by the frequency of observations. The same is true for job destruction.

The second issue is the construction of the opening and closing components. From an economic perspective, one would define a "new establishment" as a newly created institution, typically located in one place, that combines labor, capital, and purchased inputs to produce goods or services. All studies basically agree with this definition. However, because of variations across data sets in the ability to track and identify firms, studies differ in implementing this definition, which is sensitive to the treatment of mergers and acquisitions, changes in management or ownership, and the movement of establishments from one location to another.

Identifying the four employment components requires extensive data collection. At present, only three U.S. data sets are appropriate for such analyses: the Unemployment Insurance/ES202 data, the Longitudinal Research Datafile, and several extracts of Dun & Bradstreet credit records. Since all three are derived from information collected for purposes other than constructing a longitudinal file of employment, each has its strengths and weaknesses. In describing these data sets, we will concentrate on coverage, frequency of observations, firm-versus establishment-level data collection, and treatment of mergers and acquisitions.

Description of Data Sets

State-Specific Files

Unemployment Insurance (UI) tax records and ES202 reports provide state-specific data suitable for longitudinal analysis. Employers with more than a minimal number of employees (usually more than one) are required to pay taxes to finance the UI program. Because these are tax payments, states carefully monitor the filings to ensure compliance and accuracy of the returns. One drawback of the UI tax records is that they are collected at the firm level, which means that for multi-unit enterprises, data do not exist for individual plants or branches. To circumvent this problem, researchers have supplemented the UI data with ES202 records. States collect these records at the establishment level as part of the Bureau of Labor Statistics' program to enumerate employment and payroll.

By combining these two data sets, researchers have created longitudinal files of individual establishments that offer a broad coverage of industries and firms of various sizes. Data are typically collected on an annual basis so that the beginning and ending period for each interval of observations is one year. Establishments are identified by their tax identifier number, which is altered only when a significant change in corporate structure or ownership occurs. Most studies treat mergers and relocations of establishments across county boundaries as a legitimate change in an establishment's identity. Some researchers, such as Jacobson (1985), have used predecessor and successor files to track establishments more accurately and to provide a better accounting of openings and closings.

One major drawback of the UI data is their limited geographical scope. So far, information from only three states — Wisconsin, Pennsylvania, and Tennessee — has been used to study employment dynamics, although other states, including Illinois, Maryland, Ohio, and Missouri, have made their data available to researchers.

Censuses and Surveys of Manufacturers

The Census Bureau collects detailed information about manufacturing establishments on a yearly basis through the Annual Surveys of Manufacturing and on a decennial basis through the Census of Manufactures. The latter includes a com-

plete accounting of all manufacturing firms in 1967, 1972, 1977, 1982, and 1987. The annual surveys provide a probability-based sample of roughly 25 percent of these establishments.

Two different longitudinal manufacturing data sets have been constructed from the Census of Manufactures files. The first, by Dunne, Roberts, and Samuelson (1989), links the censuses, forming a panel that observes manufacturing establishments every five years. The second, constructed by the Census Bureau and called the Longitudinal Research Datafile (LRD), links both the annual surveys and the decennial censuses to form a panel with annual and quarterly observations. These data have been used by researchers, most notably Davis and Haltiwanger (1990), to estimate high-frequency employment dynamics. The primary advantage of the LRD is that it combines high-frequency observations with a sufficiently long time series to look at cyclical changes. The five-year panel used by Dunne et al., on the other hand, takes advantage of a complete census of manufacturing establishments, but misses elements of transitory or short-run employment dynamics, since establishments are created and destroyed within these five-year intervals.

The longitudinal matching of manufacturing establishments is based on plant identification, which does not change if firm mergers and acquisitions simply reflect a transfer of ownership. Although matching problems still arise (see Dunne and Roberts [1986] for details), the data set measures actual firm exits and entries as accurately as does any other source. The major drawback of the census-based files is coverage. Because these data include only manufacturing industries, they are not suitable for studying employment dynamics in other sectors and may not represent the economy as a whole.

Dun & Bradstreet Data

The Dun & Bradstreet Company maintains information on nearly 5 million businesses in every major industry and region of the country in order to assess their creditworthiness. The advantage of these data is their broad coverage of industries and regions. Birch (1981) was the first to use Dun & Bradstreet numbers to construct longitudinal files of establishments. During the early 1980s, the Small Business Administration (SBA) contracted with Catherine Armington and Marjorie Odle of the Brookings Institution to construct a

longitudinal establishment database from the Dun & Bradstreet files. We use the SBA's extract of their work later in this paper.

Data sets derived from Dun & Bradstreet files have several problems that are not present in files derived from census data. One drawback stems from the fact that the Dun & Bradstreet data set is neither a census, as is the Census of Manufactures, nor a scientifically sampled survey, such as the Survey of Manufactures. Dun & Bradstreet collects information on individual firms and establishments simply to assess their credit ratings. Therefore, biases may exist in either the identification of establishments, the number and type of establishments sampled, the frequency of sampling, or the updating of records.

In particular, Dun & Bradstreet changes an establishment's identification if it is merged or acquired. This practice may lead to overestimating the number of openings and closings, since a change in ownership is counted in both categories. Howland (1988), in examining selected industries, finds that this feature of the Dun & Bradstreet data does not induce a serious bias.

On the other hand, Dun & Bradstreet is sometimes slow to include new firms and tends to miss some openings completely, since new branches of multi-establishment firms are not counted unless they file separate credit reports. Thus, the failure to update records on a timely basis may underestimate the jobs lost due to closings and gained due to openings.^{1,2}

Jacobson (1985) compares Dun & Bradstreet data with UI data for Texas. He finds two somewhat offsetting biases. Reporting lags and failure to characterize openings and continued operations properly led the Dun & Bradstreet data to

overestimate employment and employment change from openings relative to closings in small, independent firms. At the same time, employment in large, multi-unit firms was underestimated. With these offsetting biases, Jacobson concludes that measurements of overall employment growth with Dun & Bradstreet data are reasonably accurate, but that openings may be overestimated compared to closings.

In sum, each data set has advantages and disadvantages in constructing the four employment components and in analyzing the job turnover process over time and across regions. The general consensus is that manufacturing data sets derived from census figures are probably the least problematic. However, by including only manufacturing, they provide the narrowest coverage, with only 17 percent of the U.S. workforce represented — and this share continues to decline. Thus, to provide broader coverage and the ability to generalize beyond manufacturing, it is instructive to compare employment components derived from various data sets.

II. Summary of Previous Studies

Table 1 summarizes the employment components reported by various studies that use the three data sets previously described. Comparisons among these studies are somewhat difficult: Not only do the data sets differ in construction, but wherever possible, analysts have chosen to study different years and to use intervals of different lengths in constructing the components. Even so, several similarities stand out.

First, gross employment flows are generally larger than net employment changes. For instance, Leonard (1987) finds that although net employment increased on average only 2.8 percent per year between 1977 and 1982, enough new jobs were created to boost total employment by 13.8 percent, and enough jobs were lost to reduce employment by 11 percent. While the magnitudes of these gross flows vary, all of the studies listed exhibit the same relationship between gross and net flows. Thus, net employment changes substantially understate the amount of turnover, or job creation and destruction, taking place in the market.

Leonard offers further evidence of significant job turnover not shown in the table. His analysis shows that shrinking establishments reduce their employment by an average of 21 percent per year, while growing establishments increase their employment by an average of 30 percent

■ 1 Some researchers have adjusted for this undercounting by following a two-step imputation method. First, they estimate the rate at which Dun & Bradstreet recorded start-ups between 1969 and 1980 for each of several industries. They then multiply the actual openings contained in the files by the appropriate absorption rates to approximate the incidence at which start-ups actually occurred. However, Howland (1988) and Jacobson (1985) point out several problems with this method. First, it assumes a constant absorption rate, which does not take into account the improvement in Dun & Bradstreet's recording of openings during the 12-year period. Second, it makes the unrealistic assumption that employment creation at nonsampled firms is the same as at sampled firms. Because of the company's incentive to include all active and large firms, it is more likely that unrecorded openings have fewer employees than recorded ones.

■ 2 The closing bias has been addressed in two ways. One is to assume that the establishments purged by Dun & Bradstreet are still operating and to include them in the data set. The other is to follow Dun & Bradstreet's procedure and treat the purged establishments as actual closings.

TABLE 1

Summary of Employment Components (percent)

Study	Time Period	Interval	Region	Industry	Annual Employment Change					Openings/ Creation	Closings/ Destruction	
					Openings	Expansion	Contraction	Closings	Net			
Unemployment Insurance Data												
Leonard (1987)	1977-82	1 yr.	WI	All	2.5	11.3	-8.8	-2.2	2.8	18.12	12.00	
Jacobson (1986)	1979-85	6 yr.	PA	All	5.3	2.2	-2.3	-5.0	0.1	70.67	68.49	
Dun & Bradstreet Data												
Armington and Odle (1982)	1976-82	6 yr.	U.S.	All	4.8	3.7	-2.2	-3.7	2.6	56.47	62.71	
Armington and Odle (1982)	1976-82	6 yr.	U.S.	Mfg.	3.9	3.1	-2.1	-4.0	0.9	55.71	65.57	
Eberts and Montgomery (current)	1976-78	2 yr.	U.S.	All	6.5	7.1	-4.6	-5.0	4.0	47.79	52.08	
Eberts and Montgomery (current)	1980-82	2 yr.	U.S.	All	4.3	5.6	-4.0	-5.3	0.6	43.43	56.99	
Census Bureau Data												
Dunne, Roberts, and Samuelson (1989)	1977-82	5 yr.	U.S.	Mfg.	3.5	2.3	-3.1	-3.5	-0.8	60.34	53.03	
Davis and Haltiwanger (1990)	1979-83	1 yr.	U.S.	Mfg.	1.6	6.4	-9.7	-3.0	-5.0	20.00	23.62	

NOTE: Changes are calculated as a percentage of beginning-period employment.

SOURCE: See references.

per year.³ Smaller firms tend to grow faster than larger firms, but each year a new set of small firms accounts for much of the growth. The correlation in growth rates one year apart is -0.24 , suggesting that above-average growth in one year is followed by below-average growth the next. This feature suggests that long-run growth rates may be lower than short-run changes as some firms experience frequent reversals in employment trends.

Leonard also finds substantial heterogeneity in conditions at establishments even within an industry or region (as defined by counties). In fact, there is more variation in employment growth rates *within* counties or industries than across them. The extent of this heterogeneity is reflected in the fact that the standard deviation in growth rates across establishments often exceeds the mean growth rate, especially in

manufacturing. Dunne et al. (1989) likewise find considerable heterogeneity within regions and industries. For instance, between 1977 and 1982, for every position gained in an expanding industry, 0.604 jobs were lost; for every position lost in a contracting industry, 0.644 jobs were added. Similar patterns were also found across growing and declining regions. For every position lost in a contracting region, 0.724 jobs were added, and for every position gained in an expanding region, 0.728 jobs were lost.

Second, as shown in the last two columns of table 1, there appears to be considerable variation across studies in the contribution of openings to job creation and closings to job destruction. Employment gains from openings as a share of total job creation ranges from slightly more than 18 percent to nearly 71 percent. Employment loss from firm closings as a fraction of total job destruction exhibits a similarly wide range of values. As previously discussed, the largest variations arise when intervals of different lengths are used to construct the employment components. For instance, Dunne et al.

■ 3 Weighting establishments by size and then taking the average growth rate for shrinking, growing, and stable firms yields the 2.8 percent net employment growth rate.

(1989) and Davis and Haltiwanger (1990) use virtually the same data, yet find significant differences in the contribution of openings to job creation and closings to job destruction. Dunne et al. report that 60 percent of job creation is attributable to openings, while Davis and Haltiwanger find that only 20 percent can be explained this way. The primary reason for the disparity is that Dunne et al. attribute all employment growth during the five-year interval to new firms, while Davis and Haltiwanger attribute only the first year's growth to openings, with the rest attributed to expansions. The converse applies to closings relative to contractions. Consequently, Dunne et al. find a much greater proportion of jobs created from openings or lost due to closings than do Davis and Haltiwanger.

The same large variation in employment components resulting from different observation frequencies is evident when comparing the studies of Leonard (1987) and Jacobson (1985). Both analyses use UI/ES202 data, but from different states. Therefore, the data sets are similar in construction as well as in the collection and maintenance of information (although the latter does vary across states). Yet, Leonard finds that only 18 percent of new jobs can be traced to openings when looking at observations of establishments one year apart, while Jacobson attributes 71 percent of new jobs to openings when observing establishments six years apart.

It is also worth noting that the Dun & Bradstreet and Census Bureau data yield similar results with respect to the ratio of openings to job creation. Using the Dun & Bradstreet numbers and looking only at manufacturing, Armington and Odle (1982) report that openings account for 56 percent of job creation, compared to the 60 percent found by Dunne et al. using census data. This slightly smaller fraction of jobs from openings using the Dun & Bradstreet data, even though the period was one year longer than the census-based analysis, suggests that this data set's tendency to overestimate births may not be serious. The two studies show a wider variation in the fraction of jobs lost from closings, but are still closer than studies using the same data sets but different observation frequencies.

Finally, based on the work of Armington and Odle, employment components for manufacturing closely follow employment components for all industries. The ratios of openings to job creation and closings to job destruction are quite similar, and all of the four components are reasonably close, particularly after considering manufacturing's relatively slower net employment change and, at times, employment loss.

Therefore, after accounting for differences in the intervals used to construct the employment components, it appears that the findings from various studies yield comparable qualitative results.

III. Accounting for Employment Change over Time and across Regions

To account for employment change over time and across regions, we first examine the variation of each of the four components over time in order to determine which contributes most to job fluctuations during business cycles. Similarly, we examine the variation across regions of each of the four components to identify which one is most associated with regional employment change. Some studies and data sets are more suitable for looking at one perspective than the other, but by considering evidence from the breadth of studies, a composite picture of these two processes emerges.

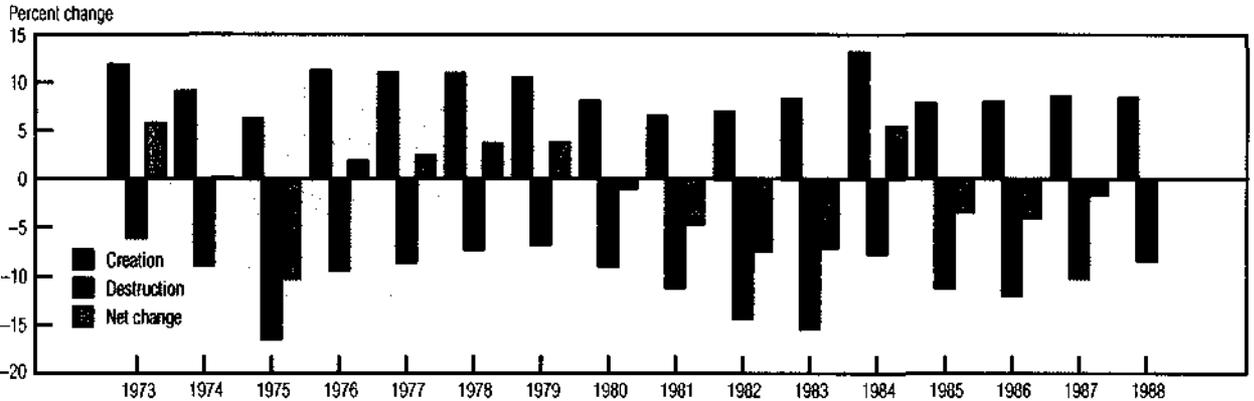
Variations over Time

Since Davis and Haltiwanger's study has the most frequent observations of the analyses discussed here, and since it spans at least two business cycles (1973–88), it is best suited for looking at the cyclical job turnover process. The results show that job destruction accounts for most of the net employment change over business cycles. As depicted in figure 1, recessions are marked by a mild decrease in creations but a large increase in destructions. Recoveries have lower-than-average destructions but slightly higher-than-average creations. The correlation between job destruction and net employment change over the period is twice as high as the correlation between job creation and net employment change (0.97 versus 0.48).

The results of Dunne et al. are consistent with those of Davis and Haltiwanger. However, because Dunne et al.'s data are not at business cycle frequencies, only tentative inferences about adjustments over these cycles can be drawn. Comparing periods of employment expansion and contraction, it appears that job destruction explains more of the variation in net employment change than does job creation. For example, the share of jobs lost from destruction rose from 19 percent in 1963–67 to 33 percent in 1967–72, as net employment fell from a 15 percent increase

FIGURE 1

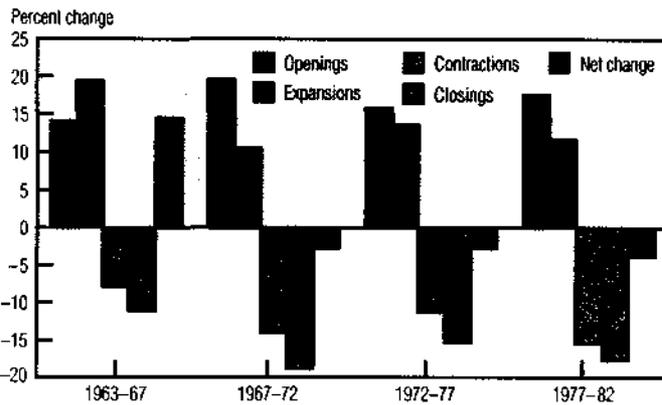
Manufacturing Job Creation and Destruction over Time



SOURCE: Census of Manufactures and Survey of Manufactures data, compiled by Davis and Haltiwanger (1990).

FIGURE 2

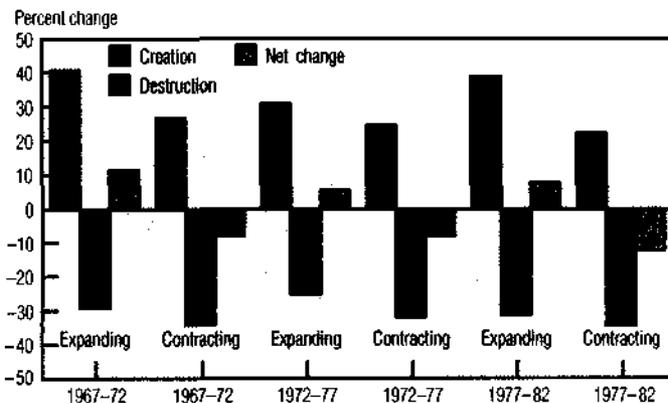
Components of Manufacturing Job Creation and Destruction



SOURCE: Dunne et al. (1989).

FIGURE 3

Manufacturing Job Creation and Destruction across Regions



SOURCE: Dunne et al. (1989).

to a 3 percent decline. For the same two periods, job gains from creations fell only moderately, from 34 to 30 percent. The same pattern emerges in comparing 1972-77 to 1977-82, as the rate of job destruction rose 6 percentage points over this interval, while the rate of job creation remained virtually unchanged.

This lack of variation in job creation reflects two offsetting trends. As seen in figure 2, job growth from expanding firms varies with net employment changes; job growth from openings runs countercyclically. Both components of job loss are procyclical and appear to be more variable than job creation components.

Leonard also offers annual time-series data, although they are much shorter than the Davis/Haltiwanger series. However, his evidence using state UI data is different from that based on census figures. Job creation is shown to be more highly correlated with net employment change than is job destruction. In addition, the variation over time of job creation is of the same magnitude as the variation of job destruction.

Variations across Regions

Dunne et al. also examine the pattern of gross flows across expanding and contracting census regions. As shown in figure 3, in two out of three cases it appears that differences in net employment change result more from variations in job creation rates than from variations in job destruction rates. During the 1967-72 period, employment gains from openings differed between the two types of regions by about 10 percentage points, while the rate of employment loss due to closings varied by less than 2 percentage points.

FIGURE 4

Manufacturing Employment Components: Expanding Regions

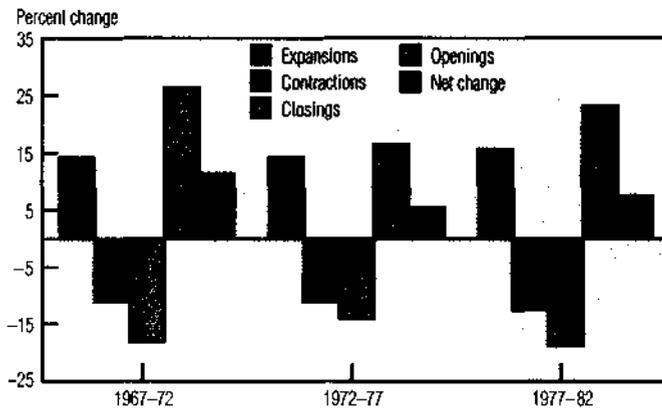
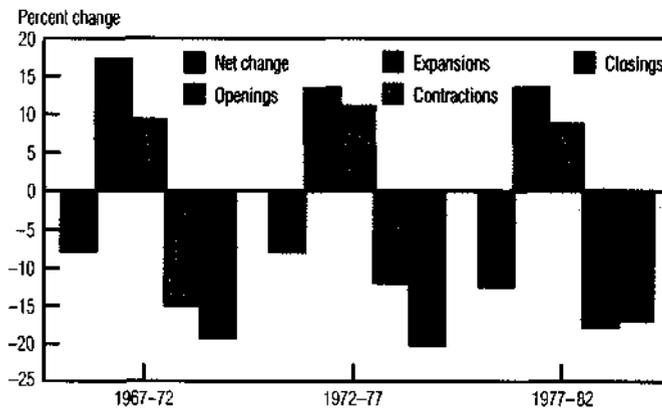


FIGURE 5

Manufacturing Employment Components: Contracting Regions



SOURCE: Dunne et al. (1989).

The same relative differences are found for the 1977-82 period.⁴

In expanding regions (figure 4), variations in the rate of openings or expansions appear to account for a larger fraction of the overall rate of net employment growth than do variations in the rate of employment loss from closings or contractions. However, for contracting regions (figure 5), changes in the rate at which employment is lost seem to be driven by variations in the rate of employment decline due to closings and contractions. This seems to suggest two different sources of manufacturing employment change. As found in the other studies, the primary source of employment variation

over time appears to be job destruction components. On the other hand, job creation, particularly from openings, appears to be the primary source of secular rates of employment change across regions. Defining regions as counties, metropolitan areas, states, or census regions does not alter the basic regional patterns of the four components of net employment change.

IV. Additional Regional Evidence

Evidence from these prior studies suggests a different pattern of gross employment flows across regions than over time. Over the business cycle (short run), job destruction behavior seems to dominate, while across regions (long run), job creation may be relatively more important. These differences need not be inconsistent any more than finding that, in the short run, aggregate demand disturbances generate most of the variations in output and yet play a minor role in explaining long-run growth differences.

The burgeoning endogenous-growth literature has focused on the factors that explain long-run growth-rate differences across countries or regions.⁵ These factors identify human capital externalities and technological spillovers (among other factors) as possible channels for the persistent differences in regional (country) growth rates. Clearly, these factors are unlikely to account for much of the short-run or cyclical variation in growth. Thus, to the extent that they are more highly correlated with job creation than with job destruction, there will be differences in the short- and long-run variability of job creation and destruction rates. In any case, a further examination of the dynamics of employment growth across regions might be useful in casting light on whether models of regional or long-run growth should focus on factors that differentially affect the job creation process.

Davis and Haltiwanger provided us with their data aggregated by census regions. We performed an analysis of variance (ANOVA) on this information to estimate the relative importance of temporal and regional variations in

⁴ The exception is the 1972-77 interval, in which employment losses resulting from closings vary more than employment gains resulting from openings. However, this period may not be representative of the nature of expanding and declining regions, as only one of the nine census regions experienced net employment losses during this time. The other two intervals offer a more balanced sample, with declining and expanding regions split evenly.

⁵ See Romer (1986), Lucas (1988), Krugman (1991), and Glaeser et al. (1992).

FIGURE 6

Variation of Manufacturing Job Creation and Destruction over Time: 1973-88

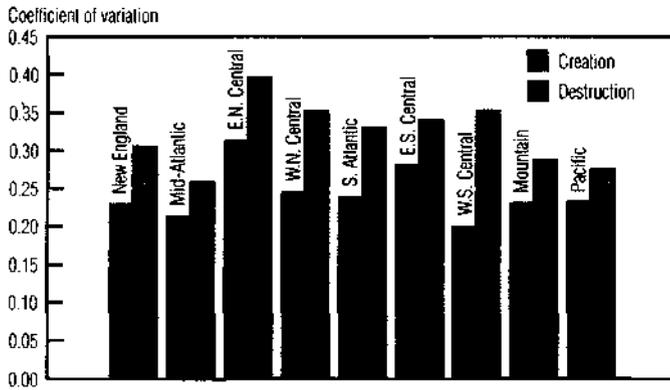


FIGURE 7

Variation of Manufacturing Job Creation and Destruction across Regions: 1973-88

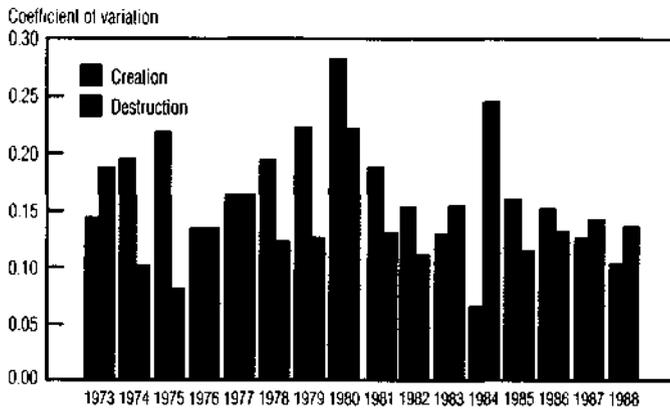
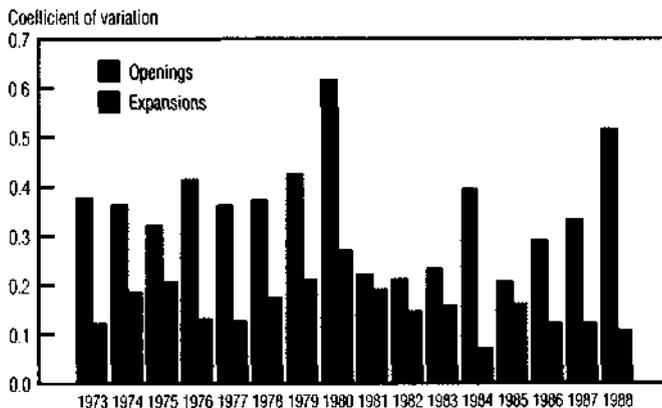


FIGURE 8

Variation of Openings and Expansions across Regions: 1973-88



SOURCE: Census of Manufactures and Survey of Manufactures data, compiled by Davis and Haltiwanger (1990).

explaining net employment change. As in most situations, the time-series variation explains a larger portion of the model variation than does regional variation. However, what is relevant for our purposes is the relative contribution of time and regional variation for job creation versus job destruction components. We found that regional variation explains a larger portion of the model variance for openings than for closings (33 percent versus 25 percent). Regional variation was also more important in explaining the model variance of expansions than of contractions (18 percent versus 3 percent).

Figure 6 presents the coefficients of variation for job creation and job destruction over time for each of the nine census regions and for all regions combined. For each region over time, job destruction varies more than job creation, which is consistent with the results for the entire sample and with the studies mentioned earlier. On the other hand, variation *across* regions is dominated by job creation (figure 7). For 11 of the 16 years covered in the sample, the variation in net employment change is explained more by fluctuations in job creation than by fluctuations in job destruction. Even during the recession years of 1981 and 1982, differences across regions in net employment change were driven principally by differences in job creation rates. The correlation across regions between net employment change and job creation is 0.69, while between net employment change and job destruction, it is 0.31.

Moreover, as illustrated in figure 8, openings vary more across regions than do expansions. However, births are not as highly correlated with net employment change as are expansions. In fact, during the 1980s, openings were primarily negatively related to regional employment conditions, with opening rates higher in the slow-growth regions. Expansions, on the other hand, are always positively related to net employment change. Therefore, Davis and Haltiwanger's manufacturing data yield the same results as do other studies: Job destruction is associated with employment change over time, while job creation is associated with employment change across regions.

To examine regional variations in job creation and destruction in more detail, we use the SBA's version of the Dun & Bradstreet data — a custom version prepared for us by SBA staff — that yields estimates of employment change due to openings, expansions, contractions, and closings for 76 industries in 263 Standard Metropolitan Statistical Areas (SMSAs).⁶ The primary

■ 6 Based on the 1977 boundary definition.

TABLE 2

**Employment Change by SMSA
(percent)**

	Overall			Expanding SMSAs			Contracting SMSAs		
	1976-78	1980-82	1984-86	1976-78	1980-82	1984-86	1976-78	1980-82	1984-86
Net change	8.0	1.6	5.9	9.6	4.6	7.4	-3.5	-3.3	-6.4
Openings	13.0	8.7	17.2	13.5	9.8	17.7	9.2	6.9	13.0
Closings	-9.9	-10.5	-14.6	-9.7	-10.6	-14.2	-11.8	-10.5	-17.4
Expansions	14.2	11.3	9.3	14.8	12.7	9.6	10.3	9.0	6.5
Contractions	-9.3	-7.9	-6.0	-9.0	-7.4	-5.7	-11.2	-8.8	-8.6
Creation	27.2	20.1	26.5	28.3	22.6	27.3	19.5	16.0	19.5
Destruction	-19.3	-18.4	-20.6	-18.7	-17.9	-19.9	-25.0	-19.3	-26.0
Gross change	46.5	38.5	47.0	47.0	40.5	47.2	42.6	35.2	45.5
Number of SMSAs	263	263	263	239	141	209	24	122	54

NOTE: Changes are calculated as a percentage of beginning-period employment. Creation is defined as openings plus expansions. Destruction is defined as closings plus contractions.

SOURCE: Authors' calculations based on the SBA's U.S. Establishment Microdata Files.

TABLE 3

**Employment Change by Expanding
and Contracting Industries (percent)**

	Overall			Expanding Industries			Contracting Industries		
	1976-78	1980-82	1984-86	1976-78	1980-82	1984-86	1976-78	1980-82	1984-86
Net change	8.0	1.6	5.9	9.6	6.9	7.9	-4.7	-4.6	-4.2
Openings	13.0	8.7	17.2	13.4	10.8	18.0	9.8	6.4	13.0
Closings	-9.9	-10.5	-14.6	-9.4	-9.8	-14.1	-13.9	-11.4	-17.1
Expansions	14.2	11.3	9.3	14.8	13.2	9.4	9.8	9.1	8.5
Contractions	-9.3	-7.9	-6.0	-9.2	-7.2	-5.5	-10.4	-8.7	-8.5
Creation	27.2	20.1	26.5	28.2	24.0	27.5	19.7	15.5	21.4
Destruction	-19.2	-18.4	-20.6	-18.6	-17.0	-19.6	-24.3	-20.1	-25.6
Gross change	46.4	38.5	47.0	46.8	41.0	47.0	44.0	35.6	47.0
Number of industries	75	75	75	61	38	55	14	37	20

NOTE: Changes are calculated as a percentage of beginning-period employment. Creation is defined as openings plus expansions. Destruction is defined as closings plus contractions.

SOURCE: Authors' calculations based on the SBA's U.S. Establishment Microdata Files.

advantages of this data set are a detailed regional breakdown and the fact that it is not limited to a single industry. Although the SBA data set is based on individual establishments, our extract of the file does not allow us access to the underlying individual firm and establishment records that stand behind our area and industry summary statistics. Thus, we cannot examine questions about within-area heterogeneity by industry.

Table 2 presents summary statistics of employment changes by source for three periods in the 1970s and 1980s. Consistent with previous studies,

we find that net employment changes substantially understate the amount of turnover in the labor market. In 1976-78 and 1984-86, gross job flows were five to eight times larger than net turnover, while in the recessionary period of 1980-82, they were more than 20 times bigger. Even if we sort SMSAs into those with declining employment and those with rising (or constant) employment, this pattern of substantially greater gross job changes than net job changes remains. Within both growing and declining regions, significant amounts of creation and destruction are

TABLE 4

Employment Change in Selected Industries (percent)

	Durable Mfg.			Nondurable Mfg.			Services			FIRE ^a		
	1976-78	1980-82	1984-86	1976-78	1980-82	1984-86	1976-78	1980-82	1984-86	1976-78	1980-82	1984-86
Net change	8.2	-5.0	-0.3	1.0	-4.8	-0.1	11.9	7.1	8.9	6.1	7.5	8.7
Openings	13.1	8.0	14.5	9.2	6.4	12.1	-12.6	8.5	17.5	12.9	6.9	16.9
Closings	-10.9	-13.1	-15.2	-10.8	-12.3	-13.7	-8.8	-8.2	-13.3	-8.1	-7.2	-12.5
Expansions	14.0	9.2	8.6	10.6	8.3	7.3	15.7	13.0	9.7	16.1	16.4	11.2
Contractions	-8.0	-9.1	-8.2	-8.0	-7.2	-5.8	-7.7	-6.3	-4.9	-14.8	-8.7	-7.0
Creation	27.1	17.1	23.1	19.8	14.8	19.4	28.3	21.5	27.2	29.0	23.3	28.1
Destruction	-18.9	-22.2	-23.4	-18.8	-19.6	-19.5	-16.5	-14.5	-18.3	-22.9	-15.9	-19.4
Gross change	46.0	39.3	46.5	38.5	34.3	38.9	44.8	36.0	45.4	51.9	39.1	47.6

a. Finance, insurance, and real estate.

NOTE: Changes are calculated as a percentage of beginning-period employment. Creation is defined as openings plus expansions. Destruction is defined as closings plus contractions.

SOURCE: Authors' calculations based on the SBA's U.S. Establishment Microdata Files.

TABLE 5

Variation in Total Employment Change

	Overall			Expanding SMSAs			Contracting SMSAs		
	1976-78	1980-82	1984-86	1976-78	1980-82	1984-86	1976-78	1980-82	1984-86
Net change	0.005	0.002	0.005	0.007	0.002	0.004	0.001	0.005	0.004
Openings	0.014	0.008	0.026	0.017	0.009	0.026	0.008	0.001	0.026
Closings	0.014	0.011	0.024	0.013	0.012	0.023	0.016	0.010	0.041
Expansions	0.016	0.014	0.009	0.019	0.016	0.009	0.010	0.009	0.005
Contractions	0.009	0.006	0.004	0.008	0.005	0.003	0.010	0.007	0.009
Creation	0.060	0.042	0.065	0.073	0.047	0.066	0.037	0.028	0.054
Destruction	0.045	0.033	0.046	0.041	0.033	0.044	0.051	0.032	0.086
Gross change	0.205	0.148	0.218	0.222	0.159	0.214	0.174	0.120	0.276

NOTE: Changes are calculated as a percentage of beginning-period employment. Variance is estimated across SMSAs. Creation is defined as openings plus expansions. Destruction is defined as closings plus contractions.

SOURCE: Authors' calculations based on the SBA's U.S. Establishment Microdata Files.

going on simultaneously. In expanding SMSAs, almost 20 percent of jobs were lost in each of our data periods, while in contracting regions, enough new jobs were created in each period to increase employment by at least 15 percent.

The same heterogeneity is displayed within industries. As shown in table 3, even contracting industries exhibit sizable employment gains from openings and expansions. For instance, while net employment in declining industries fell by 4.2 percent between 1984 and 1986, new jobs spawned from openings and expansions increased the employment base by 21.5 percent. Conversely, expanding industries are subject to significant employment losses from closings and

contractions — between 17 and 19 percent for the three periods studied. The employment change calculations for various one-digit industries, shown in table 4, reinforce the point of substantial heterogeneity within and across industries. Both the declining manufacturing industry (durables and nondurables) and the growing service and finance, insurance, and real estate industries show substantial amounts of job creation and destruction. Even in the recessionary period of 1980-82, enough manufacturing jobs were created to boost employment by 15 percent, while in the expansionary period of 1984-86, enough service-sector jobs were lost to reduce employment by 18 percent. In each sector and time

TABLE 6

Correlations of Total Employment Changes

	1976-78	1980-82	1984-86
P (Gross, Net)	0.570	0.340	0.665
P (Net, Creation)	0.875	0.799	0.899
P (Net, Destruction)	0.388	0.516	0.268
P (Gross, Creation)	0.896	0.838	0.924
P (Gross, Destruction)	-0.537	-0.630	-0.541

NOTE: Pearson correlation coefficients estimated across SMSAs.

SOURCE: Authors' calculations based on the SBA's U.S. Establishment Microdata Files.

period, gross flows were at least five times the level of net employment changes.

In table 5, we report the variance in rates of employment change between expanding and contracting SMSAs. These calculations show only a moderate amount of variation across SMSAs in rates of net employment growth — between 6 and 12 percent of the mean rate in each of our sample periods. On the other hand, the variation in gross employment changes is typically around 45 percent of the mean rate, suggesting again that there is both more turnover and more variation in turnover than would be suggested from net flow data.

These results imply that the variance across areas in openings and closings is similar to that of expansions and contractions. This finding holds in each period and in both growing and declining areas. Even if we disaggregate the data to look at employment changes by industry and SMSA, we find that openings account for more than 40 percent of creations and closings account for more than 50 percent of destructions in each period. Although similar to the results of Dunne et al., these findings differ from those of Davis and Haltiwanger, who show that openings or closings account for no more than 24 percent of job creation or destruction (see table 1). As noted above, variation in the length of the sampling intervals may explain some of this disparity. However, as others have found, change in the amount of *job creation* is the largest component of net job change across SMSAs.

In table 6, we calculate the correlation between gross and net employment flows and creation and destruction rates. In each period, job creation is more highly correlated with net job flows than is job destruction. This result is consistent with Dunne et al.'s and Davis and

Haltiwanger's finding that job creation explains a larger percentage of variations in net employment change across regions than does job destruction.

V. Conclusion

This paper offers a review and analysis of previous studies on job turnover using establishment-level data. Despite differences in the various data sets, the studies agree on several salient points. First, gross turnover is substantially greater than net growth. Second, many transitory or short-lived establishments do not show up in samples taken five years or even one year apart. Consequently, the relative contributions of openings to job creation and closings to job destruction depend on the length of the period chosen, which explains some of the differences observed across data sets. Third, substantial within-region and detailed-industry heterogeneity exists in employment growth rates.

The primary contribution of this paper is to show that the job turnover process is markedly different over time and across regions. Over time, we find that employment fluctuations are associated primarily with job destruction. Across regions, employment differences are associated more with job creation. These findings do not appear to be the result of differences in data sets, since the same data sets yield the two disparate patterns of job turnover. The results are consistent with the endogenous growth literature, which focuses on long-run factors such as human capital externalities and technological spillovers to explain long-run differences in regional or national growth rates. Since this pattern differs from the cyclical pattern of net employment dynamics, caution should be used in extrapolating models of cyclical labor market dynamics to explain long-run or regional dynamics. It will be the challenge of future research to uncover the specific factors that contribute to these differences.

References

- Armington, Catherine, and Marjorie Odle. "Small Business — How Many Jobs?" *The Brookings Review*. Winter 1982, pp. 14–17.
- Birch, David. "Who Creates Jobs?" *Public Interest*, vol. 65 (Fall 1981), pp. 3–14.
- Davis, Stephen, and John Haltiwanger. "Gross Job Creation and Destruction: Microeconomic Evidence and Macroeconomic Implications." *NBER Macroeconomics Annual 1990*, vol. 5, pp. 123–68.
- Dunne, Timothy, and Mark J. Roberts. "Measuring Firm Entry, Growth, and Exit with Census of Manufactures Data." Pennsylvania State University, unpublished manuscript, 1986.
- _____, _____, and Larry Samuelson. "The Growth and Failure of U.S. Manufacturing Plants," *Quarterly Journal of Economics*, vol. 104, no. 4 (November 1989), pp. 671–98.
- Glaeser, Edward L., et al. "Growth in Cities," *Journal of Political Economy*, vol. 100, no. 6 (December 1992), pp. 1126–52.
- Howland, Marie. *Plant Closings and Worker Displacement: The Regional Issues*. Kalamazoo, Mich.: W.E. Upjohn Institute for Employment Research, 1988.
- Jacobson, Louis. "Analysis of the Accuracy of SBA's Small Business Data Base," Center for Naval Analyses, Working Paper 85-1958, Alexandria, Va., October 23, 1985.
- _____. "Job Creation and Destruction in Pennsylvania, 1975–85," W.E. Upjohn Institute for Employment Research, unpublished manuscript, 1986.
- Krugman, Paul. "Increasing Returns and Economic Geography," *Journal of Political Economy*, vol. 99, no. 3 (June 1991), pp. 483–99.
- Leonard, Jonathan. "In the Wrong Place at the Wrong Time: The Extent of Frictional and Structural Unemployment," in Kevin Lang and Jonathan Leonard, eds., *Unemployment and the Structure of Labor Markets*. New York: Basil Blackwell, 1987.
- Lucas, Robert E., Jr. "On the Mechanics of Economic Development," *Journal of Monetary Economics*, vol. 22, no. 1 (July 1988), pp. 3–42.
- Romer, Paul. "Increasing Returns and Long-Run Growth," *Journal of Political Economy*, vol. 94, no. 5 (October 1986), pp. 1002–37.

A Monte Carlo Examination of Bias Tests in Mortgage Lending

by Paul W. Bauer and Brian A. Cromwell

Paul W. Bauer is an economist at the Federal Reserve Bank of Cleveland, and Brian A. Cromwell is a manager at Deloitte & Touche, Washington, D.C. The authors thank Robert Avery, Patricia Beeson, Paul Calem, Fred Furlong, Arden Hall, Gordon Smith, Mark Sniderman, and David Vandre for useful discussions and suggestions. Karen Deangelis provided excellent research assistance. The authors also thank the Federal Reserve Bank of San Francisco for supporting this research while Mr. Cromwell was an economist there.

Introduction

For three years, data released through the Home Mortgage Disclosure Act (HMDA) have documented that blacks are denied loans at a much higher rate than whites.¹ Whether this differential reveals bias by lending institutions, however, is a hotly debated issue. One reason is that HMDA data do not include all of the information contained on loan applications, such as the applicants' job and credit histories and the size of their down payments. Using information from loan applications, the Federal Reserve Bank of Boston (Munnell et al. [1992]) conducted a large statistical study of Boston-area banks (3,062 mortgage applications). The authors found that although the gap in denial rates between blacks and whites narrows when the additional information from the loan file is included, a statistically significant differential remains. Far from settling the issue, however, the Boston Fed study has merely provided the basis for further analysis. Interestingly, the Federal Deposit Insurance Corporation

(FDIC) examined the loan applications identified as most likely to have been rejected because of bias, but found no evidence of bias and raised a number of methodological and empirical problems with the Boston Fed study (Home [1994]).

This *Economic Review* explores the effectiveness of testing procedures in uncovering discrimination by mortgage lenders. After outlining the regulatory issues and the inherent problems in testing for bias, we investigate how well various tests perform under a variety of circumstances using simulated mortgage loan applications. We discuss the problems involved in testing real-world data and demonstrate how these tests perform using artificially generated data in which we can control the degree of bias. It should be emphasized that we cannot answer the question of whether there is bias in mortgage lending, but we have a great deal to say about the performance of bias tests when there are no measurement problems with the data (a condition for which all researchers strive) and when the degree of bias is known beforehand.

We find that tests employing all of the information included in our simulated loan files perform much better than those using only the HMDA subset of data. This result is not unexpected, yet the

■ 1 Under HMDA, lending institutions are required to record and report data on applicants' race, sex, income, type of loan, loan amount, and whether the loan was approved or denied.

B O X 1

Legal Definitions of Discrimination and Current Testing Methods

Legal Definitions

Several regulatory agencies are responsible for ensuring that discrimination does not occur in lending institutions. Each agency has developed a means of testing for discrimination in the bank class it is responsible for overseeing. Three general types of discrimination are recognized: overt discrimination, disparate treatment, and disparate impact.^a

Overt discrimination occurs "when a lender openly discriminates on a prohibited basis." In addition, overt discrimination exists even when a lender expresses, but does not act on, a discriminatory preference. For example, a lender may offer two equally qualified applicants of different races different credit limits. Regulatory agencies would classify this action as overt discrimination.

There is evidence of *disparate treatment* when "a lender treats applicants differently based on prohibited factors." Disparate treatment may range from overt discrimination to subtle differences in treatment. For example, a lender may provide a nonminority applicant with more assistance in the application process than it would a minority applicant.

Disparate impact occurs when a lender "applies a practice uniformly to all applicants but the practice has a discriminatory effect on a prohibited basis and is not justified by business necessity." For example, a high-minimum-loan requirement may prevent low-income housing applicants, who are typically minorities, from being granted a loan.

Testing Methods

In order to test for lending discrimination, three procedures have been developed: testing, matched pairs, and statistical analyses. Each regulatory agency's method utilizes one or a combination of these general procedures.

Testing is a means of measuring differences in treatment among loan applicants. It involves sending "testers" disguised as loan applicants into an institution where they attempt to apply for a loan. Treatment of the "applicants" is then compared by the regulatory agency and determinations are made concerning the existence of discrimination.

Matched pairs are a means of grouping minority and nonminority applicants in a manner that will allow for accurate comparison of treatment between groups. Matched pairs are determined by comparing loan-to-value and debt-to-income ratios among applicants. Matched pairs are compared in a manner similar to that used for testers.

Statistical analyses may also be used to test for discrimination. An institution's lending data from previous years are collected, entered into a statistical program, and then analyzed for evidence of discrimination.

a. These definitions are outlined in "Interagency Task Force on Fair Lending," *Federal Register*, vol. 59, no. 73 (April 15, 1994), p. 18268.

difficulty of detecting low levels of bias even with large sample sizes is somewhat surprising. With the bias parameter set to the level that results in rejection rates similar to the actual HMDA data, sample size is crucial. At this level of bias, tests with sample sizes under 50 almost always fail to detect bias, whereas econometric tests with sample sizes greater than 200 perform well. Finally, we conduct nonparametric tests that have their roots in the procedures employed by examiners. Although these tests work very well in small samples, they also tend to find bias even in simulations when it is not present.

I. Methods of Testing for Bias

The Equal Credit Opportunity Act prohibits discrimination with respect to any aspect of a credit transaction based on race, color, religion, national origin, sex, marital status, age (provided the applicant has the capacity to contract), receipt of income from public assistance programs, and good-faith exercise of any rights under the Consumer Credit Protection Act. Regulatory institutions such as the FDIC, the Office of the Comptroller of the Currency, and the Federal Reserve Board are charged with enforcing this Act and uncovering discriminatory credit approval processes (see box 1). Of the thousands of consumer examinations conducted each year, few indicate credit discrimination on the basis of race.² In 1992, about 90 percent of the 5,602 banking institutions in the United States received outstanding or satisfactory ratings on their consumer exams.

On the other hand, over the last three years, large disparities between credit approval rates of white and minority applicants have been revealed by the revised HMDA data. Even controlling for income, minority applicants were rejected for credit at rates two to three times those of white applicants. This result potentially indicates widespread discrimination on the basis of race. A competing explanation for this credit-approval disparity, though, is that minority populations are commonly found to be less creditworthy (for example, because of lower asset levels) than the nonminority population. The revised HMDA data, however, do not include relevant financial information on credit

■ 2 A consumer exam is conducted to ensure that the regulated financial institution is in compliance with the various statutes relating to the treatment of consumers, such as the Equal Credit Opportunity Act and the Community Reinvestment Act.

applicants (such as assets and credit history) that is available to decisionmaking institutions.³

The issue of bias in mortgage lending is a broad one, and some researchers have raised the concern that simple comparisons of lenders' denial rates are not sufficient for grasping the complexities surrounding community-oriented lending.⁴ Our purpose here is to explore the narrower issue of looking at the performance of tests that examiners could use to detect bias in the course of their regulatory duties.

Examiners have access to the complete loan files for both approved and rejected credit applications and consequently are able to look at the financial information missing from the HMDA data. In the past, applicant profiles were constructed for a sample of white and minority applicants (both acceptances and rejections). No formal statistical test was conducted, but the examiner looked for evidence that applicants are treated according to the articulated lending criteria of the institution. Financial institutions are required to maintain such criteria, and inspection of them is part of the exam process.

Currently, the Federal Reserve is implementing a testing procedure that estimates a logit model using the HMDA data for an institution over roughly the previous three years. If the race variable is found to be significant, then a random sample of loan files is selected and the model is reestimated after adding pertinent non-HMDA variables, such as employment and credit histories, net worth, and amount of other debt obtained from the loan application. If the race variable continues to be significant, the examiners pull the loan applications that the estimated model predicts were influenced by bias and seek out the bank's management for an explanation.

The sample size of applications actually examined, however, is constrained by the time (and number) of examiners that the agency is able to devote to the procedure. Fed guidelines suggest constructing a matched sample of

100 white and 100 minority applications. Constrained regulatory resources thus potentially undermine the effectiveness of a consumer exam in uncovering bias. Clearly, two important questions are whether pretests employing the HMDA data (which are relatively costless to the examiners, but not to the lenders) provide useful information, and how large a sample is required to determine whether lending bias exists. After developing a simulation model that allows us to vary the amount of bias against minorities, we use it to see how well various testing procedures can identify an institution that discriminates.

II. Simulation Model

Before going into the details of our simulation method, a brief overview is useful to highlight the key parts of the Monte Carlo process.

Throughout the discussion of the model and consequently of our findings, it must be remembered that this is a simulation model and thus cannot answer the question of whether lending institutions are really subject to bias. Our goal is to explore how well various tests for bias perform when the level of bias is known beforehand. To accomplish this task, it is not necessary to mimic the underlying real-world process precisely. The key qualitative characteristics we wish to simulate are 1) that lenders base their mortgage approval decision on a larger set of variables than is included in the HMDA data, 2) that some of these omitted variables are correlated with race, and 3) that we can control the degree to which our simulated lender allows race to influence the loan approval process.

The first step is to generate a pool of loan applicants to simulate the actual population of both nonminorities and minorities in terms of income, net worth, debt payments, and credit history. Wherever possible, the variables are calibrated using the results of actual consumer surveys. These generated applicants then apply for loans in a credit approval model that is representative of actual approval processes used by financial institutions. The credit approval model allows for the possibility of bias against minorities, with the level of discrimination able to be varied from zero (in which case credit decisions are made solely on the basis of financial characteristics) to a level that results in a significantly higher level of rejections for minority applications. The results from the credit approval model are a set of loan files

■ **3** Avery, Beeson, and Sniderman (1993) cite results from an extremely large regression on the national HMDA data set, controlling for institutional and neighborhood characteristics and available individual information, and find a 7 to 10 percent unexplained differential linked to race. Munnell et al. (1992) explore the importance of the missing financial information in evaluating lending decisions in the Boston metropolitan statistical area. They find that differences in financial characteristics explain 9.9 percentage points of the observed 17.8 percentage-point discrepancy in denial rates of whites and minorities. The remaining 7.9 percentage points are considered to be linked to race. A discussion of a number of systematic problems present in HMDA data can be found in Horne (1994).

■ **4** See Avery, Beeson, and Sniderman (1993).

with applicant information and a 0/1 variable indicating whether or not the loan was granted. Although we have tried to benchmark our generated applicants to nationally reported data, this was not possible in all cases. Our numerical results will be sensitive to changes in the applicant generation process, but the qualitative import of our results will not.

At this point, our simulated examiner extracts a sample from the set of loan files and tests for discrimination. Several tests are possible. A "bank examiner" approach could search for evidence that whites and minorities with similar characteristics are treated differently, perhaps through matching rejected minority applications with approved white applications. Various levels of sophistication are possible. Alternatively, an "econometric" approach would estimate an equation and test for a significant coefficient on the variable representing race. In either the "examiner" or "econometric" approach, we will take repeated draws from the loan-file population and measure the proportion of times the test indicates a positive result for discrimination. By running the tests on loan files generated from a discriminatory credit approval process, we are able to explore the sensitivity of various tests for discrimination.

Generation of Applicant Data

The applicant sample is generated with the following characteristics: income, net worth, loan amount, other debt payments, credit history, and race. Actual loan applications would contain many more variables, but in our model these are the only ones the bank considers. More variables could be incorporated into the simulation model, but their addition would be unlikely to alter the basic thrust of our findings. Where possible, we have initially calibrated the means and correlations of these variables to those from consumer financial surveys and other sources.⁵

To generate the samples, we first created a matrix of the variances and covariances of the financial variables for the white and minority populations. The covariances of the loan amount and income (in log form) were identified from the 1990 national set of HMDA data for both nonminority and minority populations. We do

not have information on the correlation of loan amount and the other financial variables, so we set these to plausible values. The means of the sample for income and loan amount were also determined using the HMDA data. We set the means for other financial variables using information from the 1989 Survey of Consumer Finance (SCF), a nationally representative wealth survey.⁶ In particular, the mean of net assets in the sample was established by multiplying the mean income in the HMDA data by the ratio of assets to income in the SCF (for white and minority populations, respectively). The variance of assets and the correlation between assets and income for white and minority populations were also derived from this survey. We determined the mean of "other debt" payments using the ratio of other debt payments to income.

Our information on credit history for real loan applicants is limited. We used the answers to the SCF question on the timely loan and credit card payments to establish the sign of the correlation between bad credit history and the other financial variables, and modeled the tendency to have credit problems as an underlying normal random variable (larger values of credit history are considered bad) that correlates negatively with income and net worth.

Given the means, variances, and correlations, the applicant sample was generated by 1) multiplying the draws from the log-normal distribution by the Cholesky decomposition of the desired covariance matrix, and 2) rescaling the resulting series to match the desired means. This procedure ensures that the generated sample exhibits the desired correlations across variables. Credit history is rescaled into a categorical variable as follows: For whites, about 5 percent will have serious credit problems, 25 percent minor problems, and 70 percent no credit problems. For minorities, the corresponding percentages are roughly 7, 31, and 62, so that by construction they have a higher incidence of credit problems. These thresholds are arbitrarily chosen to give minorities more credit history problems in order to match the qualitative characteristics of real-world data.

The people surveyed in the SCF do not necessarily represent the population of potential mortgage applicants, because potential homeowners tend to be more affluent than the population as a whole. For our initial set of

■ 5 We focus on race in this paper, but a similar approach could be used to determine the power of tests for lending bias related to sex, age, and marital status.

■ 6 See Kennickell and Shack-Marquez (1992) for further information on the SCF.

TABLE 1

**Sample Means of Generated Population
(Sample size of 5,000, 50% minority)**

	White	Minority
Income (annual)	\$63,728	\$36,029
Net worth	\$291,682	\$36,455
Loan amount	\$73,744	\$45,561
Loan payment (monthly)	\$647	\$400
Other debt payment (monthly)	\$452	\$202
Minor credit problems (percent)	25	31
Major credit problems (percent)	5	7

SOURCE: Authors' calculations.

variables, however, we found that rejection rates for both whites and blacks from the credit model (presented in the next section) were "too high" in comparison to those found in actual HMDA data, due in part to the low level of assets of minority families seen in the SCF. To adjust for this, we marginally increased the income and net assets of minority families, and marginally reduced the loan amount for both blacks and whites. The resulting generated samples should be viewed as broadly representative of the financial characteristics seen in the actual white and minority population, but as only partially calibrated due to lack of information on the financial characteristics of mortgage applicants.⁷ Changing the financial characteristics, of course, does affect the probability of acceptance or rejection, but is unlikely to change the qualitative characteristics of our results.

The sample means of a draw of 5,000 applicants (half minorities) from our samples are reported in table 1.⁸ Corresponding sample correlations are reported in table 2. Our white applicants (for this draw) have significantly larger incomes and net worths than do minorities, consistent with SCF data. Correspondingly, average loan amounts are higher for whites than for minorities. Our sample was generated so that positive correlations would be observed between income, net worth, loan amount, and other debt, and a negative correlation seen between finan-

cial variables and credit history. Again, we view this generated sample as only partially (and imprecisely) calibrated, but as reflecting broad relationships observed in the financial characteristics of populations in the real world.

Credit Approval Model

Once our applicant pool is generated, the "forms" are fed into our credit approval model that determines whether or not the financial institution makes the loan. The process is modeled so that "good" applications are almost always approved, and "bad" applications are almost always rejected. Borderline applications are approved or denied with a probability determined by the number of problems in the application, and by race in the case of a discriminatory bank.

We assume that the application is initially for a 30-year loan at a 10 percent interest rate with monthly payments and a 20 percent down payment.⁹ The loan amount is initially determined through the applicant generation model. However, an applicant is unlikely to apply for a 20 percent down payment loan if he lacks the necessary assets. We model this down payment decision process in the following way: First, if the 20 percent down payment is greater than the applicant's net worth (plus two monthly payments), the applicant shifts to a 10 percent down payment. The loan amount and monthly payments are recalculated accordingly. Second, if net worth still falls short of the down payment, the applicant shifts to a loan with a 5 percent down payment.¹⁰ Setting the loan amount and down payment in this sequential fashion is somewhat arbitrary, but it allows marginal applicants to apply for appropriate loans. Imposing a strong positive correlation between loan amount and net worth further tends to prevent paupers from applying for million-dollar mortgages.

Loan applications are scored according to four standard criteria: 1) the ratio of loan payment to income, 2) the ratio of total debt payment to income, 3) the percentage of the down payment, and 4) credit history. Any of the first three criteria can result in automatic rejection if

■ 9 Varying the interest rate and the term of the loan would introduce diversionary complications into the simulation model.

■ 10 Not addressed in this version of the model is the decision of the private mortgage insurer for down payments below 20 percent, or the effect of government insurance programs on loans with 5 percent down payments. Again, these factors are unlikely to affect the basic thrust of our results.

■ 7 Supplementing the SCF data with information in Munnell et al (1992) is one possible strategy for correcting this shortcoming.

■ 8 Our reason for oversampling is discussed later.

TABLE 2

**Sample Correlations of Generated Population
(Sample size of 6,000, 50% minority)**

	Income (Annual)	Net Worth	Loan Amount	Other Debt	Minor Credit Problems	Major Credit Problems	White
Income (annual)	1.000	0.341	0.354	0.076	-0.406	-0.237	0.349
Net worth	0.341	1.000	0.336	0.048	-0.078	-0.047	0.180
Loan amount	0.354	0.336	1.000	0.095	-0.170	-0.123	0.247
Other debt payment	0.076	0.048	0.095	1.000	-0.009	-0.020	0.301
Minor credit problems	-0.406	-0.078	-0.170	-0.009	1.000	-0.156	-0.069
Major credit problems	-0.237	-0.047	-0.123	-0.020	-0.156	1.000	-0.030
White	0.349	0.180	0.247	0.301	-0.069	-0.030	1.000

SOURCE: Authors' calculations.

it is violated. Each of the criteria also has a "borderline" gray area (called GRAY1, GRAY2, GRAY3, and GRAY4, respectively) that results in a positive probability of rejection. Since the fourth criterion is a qualitative variable, possibly subject to varying interpretations of its severity, an "autoreject" here means that failing this criterion, by itself, results in a 50 percent chance of denial.¹¹ If all four criteria meet approval, then the application is almost always automatically accepted. With real loan applications, several other criteria (such as employment history and an appraisal) are considered during the underwriting process, but we focus on just these four standard criteria in an attempt to make our model more tractable.

The regions for the four criteria are as follows:

- 1) **Loan payment to income (PMT/Y):**
 If $PMT/Y > 0.40$, then reject the application;
 if $0.40 > PMT/Y > 0.28$, then GRAY1;
 if $PMT/Y < 0.28$, then the application passes this criterion.
- 2) **Total debt payment to income (TPMT/Y):**
 If $TPMT/Y > 0.48$, then reject the application;
 if $0.48 > TPMT/Y > 0.36$, then GRAY2;
 if $TPMT/Y < 0.36$, then the application passes this criterion.
- 3) **Net worth (NW):**
 If $\text{down payment} + 2 \times PMT < \text{net worth}$,
 then reject the application;
 if 5 or 10 percent down payment, then GRAY3;

if 20 percent down payment, then the application passes this criterion.

4) Credit history:

If there are major credit problems, then randomly reject the application half the time;
 if there are minor credit problems, then GRAY4;
 if there are no credit problems, then the application passes this criterion.

These credit rules are motivated by actual credit processes. The financial ratios of 28 and 36 percent in rules 1 and 2, respectively, mirror actual tests used by financial institutions and suggested by secondary market purchasers, such as the Federal National Mortgage Association (FNMA).¹² Rule 3 checks for the level of down payment and a minimum net worth. Rule 4 seeks evidence of major and minor credit problems.

We allow for gray areas, however, in order to mimic the judgment that goes into the credit process for borderline applications. For example, the financial ratios suggested by FNMA are considered guidelines subject to the discretion of the lender. The down payment requirement reflects the bank's adjustment for an increased likelihood of default on low-down-payment loans. Finally, allowing for major and minor

■ 11 An additional consideration is that logit regressions cannot handle an independent variable that is too highly correlated with the dependent variable.

■ 12 See Federal National Mortgage Association (1992), pp. 601-94.

TABLE 3

Loan Scoring of Generated Population
(Percent of sample, sample size
of 5,000, 50% minority)

	White	Minority
AUTO APPROVE (Meets all criteria)	65	49
BORDERLINE (Violates some criteria)		
GRAY1 Payment/income between 28 and 40 percent	6	7
GRAY2 (Payment + other debt)/income between 36 and 48 percent	9	9
GRAY3 Down payment below 20 percent	6	26
GRAY4 Minor credit problems	25	31
AUTOREJECTS (Serious problems)		
AUTOR1 Payment/income above 40 percent	6	8
AUTOR2 (Payment + other debt)/income above 48 percent	14	13
AUTOR3 Net worth below down payment plus 2 PMTs	3	15
AUTOR4 Major credit problems	5	7

SOURCE: Authors' calculations.

credit problems allows for the distinction between recent bankruptcies versus a couple of late payments. Past credit problems may also be the result of unusual circumstances. The more GRAY areas that an application hits, however, the more likely that it will be rejected. In addition, we include a small probability of rejection of "clean" applications with no GRAYs to reflect some randomness in the decision process. The probability of approval contingent on the total number of GRAY areas is modeled as follows:

If TOTAL GRAYs = 0, then 3 percent rejection rate;

If TOTAL GRAYs = 1, then 20 percent rejection rate;

If TOTAL GRAYs = 2, then 30 percent rejection rate;

If TOTAL GRAYs = 3, then 40 percent rejection rate; and

If TOTAL GRAYs = 4, then 50 percent rejection rate.

These rates were chosen in order to generate a plausible number of rejections corresponding to the severity of credit problems.

We also use this process for borderline applications to introduce discrimination against minority applicants. Given this modeling, discrimination occurs because minorities are more likely than nonminorities to be turned down for a loan when there are blemishes in their loan applications. In general, we multiply the vector of approval probabilities by a bias parameter (BIAS) to increase the probability of rejection of minority applications.

If TOTAL GRAYs = 0, then $1/(1 - \text{BIAS})$ percent rejection rate;

If TOTAL GRAYs = 1, then $20/(1 - \text{BIAS})$ percent rejection rate;

If TOTAL GRAYs = 2, then $30/(1 - \text{BIAS})$ percent rejection rate;

If TOTAL GRAYs = 3, then $40/(1 - \text{BIAS})$ percent rejection rate;

If TOTAL GRAYs = 4, then $50/(1 - \text{BIAS})$ percent rejection rate.

There are many ways to introduce bias into the loan approval process. This approach has the advantage of employing only a single parameter that can easily be varied from no bias (BIAS = 0) to the point where no minorities ever receive loans (BIAS = 1).

For example, if the bias parameter is set to 0.5, then minority applicants with a single GRAY will be rejected 40 percent of the time, applicants with two GRAYs will be rejected 60 percent of the time, and applicants with three or four GRAYs will always be rejected at 80 and 100 percent rates, respectively. We use this simple model so that we can easily test (by varying one parameter) the sensitivity of the results to varying levels of discrimination. Although more complicated models of discrimination can be used, we believe this model adequately captures the flavor of a discriminatory process where minority applicants are less likely to be approved in borderline cases.¹³

The sample statistics for GRAY1 - GRAY4 for whites and minorities are shown in table 3 (the sample includes 2,500 whites and 2,500 minorities).

■ 13 This result is implied by the findings of Munnell et al. (1992).

TABLE 4

**Credit Application Decisions
(5,000 draws, 50% minority)**

	Discrimination Parameter ^a				
	0.0	0.2	0.4	0.6	0.8
Percent approved					
Total	71.3	70.7	67.1	59.0	54.0
White	76.1	76.1	76.1	76.1	76.1
Minority	66.4	65.2	58.0	41.9	31.9
Percent denied					
Total	28.7	29.3	32.9	41.0	46.0
White	23.9	23.9	23.9	23.9	23.9
Minority	33.6	34.8	42.0	58.1	68.1
Minority/white					
Percentage point difference	9.7	10.9	18.1	34.2	44.2
Due to financial characteristics	9.7	9.7	9.7	9.7	9.7
Due to discrimination	0.0	1.2	8.4	24.5	34.5

a. If zero, no bias. If one, no loans made to minorities.
SOURCE: Authors' calculations.

The proportion automatically rejected is also shown along with the four reasons for rejection, AUTOR1 – AUTOR4. (Applicants may have multiple reasons for automatic rejection.) By construction, whites are more likely than minorities to have clean applications that are automatically approved and are less likely to be automatically rejected. In our model, the most common GRAY areas hit are the credit history rule for both races and the down payment rule for minorities. Forty percent of the loans to minorities are for down payments below 20 percent, and of these, 15 percent are still automatically rejected for not having the necessary net worth for a mortgage with a 5 percent down payment. In addition, 13 percent are rejected for a high ratio of total debt to income. For whites, this financial ratio is also the most common reason for automatic rejection (14 percent).

Approval rates for varying levels of bias are shown in table 4. Whites are approved about 76.1 percent of the time in our model. With no discrimination, the minority approval rate is 66.4 percent, a difference of 9.7 percent. This compares with a national approval rate of 75.5

percent for whites and 55.7 percent for minorities (a 19.8 percentage-point difference) observed in 1990 national HMDA statistics. In our model, this difference is due solely to the financial characteristics of the applicant, and it is also close to that attributed to financial characteristics in Munnell et al. (1992). By varying our BIAS parameter, however, we can generate approval rates for minorities that mimic the observed approval rates in the HMDA data.

A small level of bias (0.2) results in only a slight increase in the disparity between whites and minorities. The next level of bias reported (0.4) raises the disparity to 18.1 percentage points, close to the observed 19.8 percentage-point difference in the national statistics. Raising the level of bias to 0.8 results in a rejection rate for minorities that is almost three times that of whites. Thus, our credit approval/discrimination model can generate the range of credit approvals observed in the HMDA data, allows for easy variation of the level of discrimination, and generates loan files that can be used to test for discrimination through a bank examination.

TABLE 5

**Proportion Passing Examination
(1,000 repetitions, logit test)**

Sample size	Discrimination Parameter—Full Set of Variables ^a				
	0.0	0.2	0.4	0.6	0.8
50	0.990	0.981	0.976	0.875	0.555
75	0.970	0.932	0.892	0.701	0.238
100	0.971	0.937	0.856	0.572	0.068
150	0.971	0.929	0.824	0.441	0.005
200	0.971	0.936	0.777	0.323	0.003
250	0.972	0.925	0.720	0.184	0.000
300	0.971	0.923	0.695	0.161	0.000
350	0.966	0.914	0.656	0.103	0.000
400	0.970	0.902	0.584	0.062	0.000
450	0.974	0.908	0.568	0.040	0.000
500	0.967	0.883	0.505	0.027	0.000
600	0.969	0.870	0.456	0.019	0.000
800	0.973	0.850	0.339	0.001	0.000

Sample size	Discrimination Parameter—HMDA Variables ^a				
	0.0	0.2	0.4	0.6	0.8
50	0.926	0.905	0.835	0.693	0.265
75	0.876	0.845	0.734	0.497	0.112
100	0.879	0.790	0.676	0.406	0.038
150	0.828	0.728	0.527	0.203	0.007
200	0.760	0.671	0.412	0.109	0.000
250	0.744	0.564	0.341	0.040	0.000
300	0.648	0.506	0.272	0.027	0.000
350	0.635	0.446	0.182	0.015	0.000
400	0.586	0.383	0.136	0.005	0.000
450	0.572	0.333	0.110	0.003	0.000
500	0.521	0.291	0.068	0.001	0.000
600	0.423	0.214	0.041	0.000	0.000
800	0.316	0.109	0.016	0.000	0.000

a. If zero, no bias. If one, no loans made to minorities.
SOURCE: Authors' calculations.

III. Analysis of Econometric Tests

In this section, we test the statistical power of econometric exam tools through Monte Carlo simulation. We vary both the sample size and level of bias to test the sensitivity of the exams to these factors. Repeated draws (of our preset sample size) from bank loan files (with our preset level of bias) are used in a logit regression. The dependent variable is a 0/1 variable indicating credit approval. Independent variables include those corresponding to the credit approval process: income, net worth, payment/income, total

debt payments/income, down payment/net worth, CREDIT1 (0/1 dummy for minor credit problems), and CREDIT2 (0/1 dummy for major credit problems). In addition, a 0/1 dummy variable indicating a minority is included. A significant (negative) estimated coefficient is taken as a positive test for discrimination.

Logit is the preferred estimator on theoretical grounds for such a model because it allows for the 0/1 nature of the dependent variable (determining whether the loan was approved) and for slightly more outliers than the Probit model.¹⁴ These advantages have also led to its use in the Boston Fed study and in the current Federal Reserve testing procedure for investigating possible lending bias. We conducted 1,000 repetitions for each setting of the model, oversampling minorities so that they compose 50 percent of the sample.¹⁵

Table 5 reports the proportion of examinations that "passed" (failed to find statistically significant evidence of discrimination) at the standard 5 percent significance level. Figure 1 plots the same data, but is useful for illustrating how the performance of the test improves as the sample size increases. In the first column, there is no bias, yet some banks fail to pass. The level of "false positives" is an important factor in evaluating the usefulness of a test. False positives represent the risk of erroneously accusing a bank of discriminating when it in fact does not. An ideal test would always find bias when it is present, but would never find it when it is absent. For logit, this rate is typically in the 1 to 3 percent range over the sample sizes studied and tends to be slightly better than the 5 percent we allowed for in our selection of the significance level.

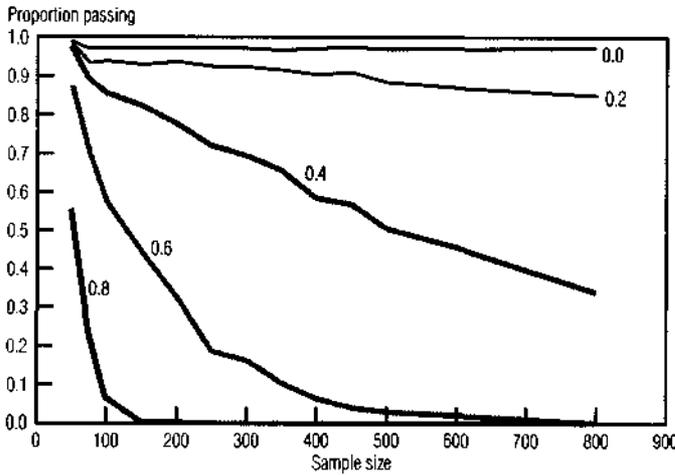
In the second column, we set our bias parameter to 0.2, introducing a low-level bias that, as seen in table 4, increases the rejection rate for minorities only slightly. For small sample sizes, we rarely find evidence of this discrimination. Even for a sample size of 800, our tests successfully uncover discrimination only 15 percent of the time. A small level of discrimination can go undetected by statistical methods even with very large samples.

■ 14 We know that banks do not use an equation like this to make their decisions, but we use it in our model to approximate their decision-making process.

■ 15 In earlier work, we explored the importance of oversampling. We found that it increases the statistical power of the exam by a very small amount, but reduces the incidence of false positives in small sample sizes. Oversampling also avoids anomalous results in small sample sizes, such as having no minority acceptances.

FIGURE 1

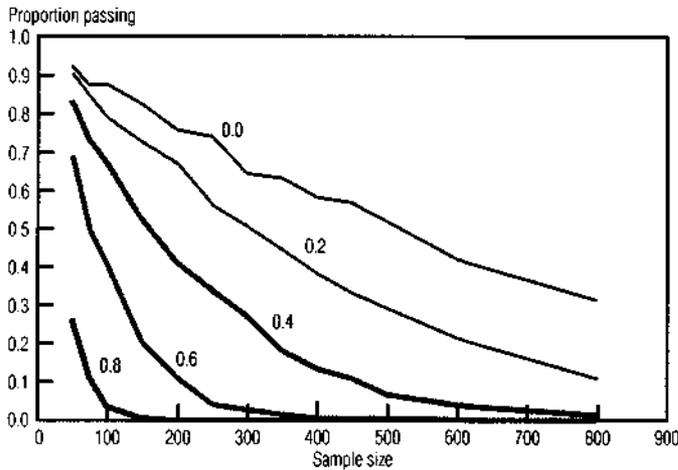
Power of Logit Test— Full Variable Set



SOURCE: Authors' calculations.

FIGURE 2

Power of Logit Test— HMDA Variables



SOURCE: Authors' calculations.

At moderate levels of discrimination (with bias equal to 0.4, raising the lending differential between whites and minorities significantly, as seen in table 4), sample size plays a critical role in the power of the exam. We find discrimination less than 15 percent of the time for sample sizes of 100 or less. Raising the sample size from 100 to 200 increases the power of the exam to about 23 percent, and a sample size of 800 results in a statistical power of 66 percent.

At larger levels of bias, the logit test is better able to detect discrimination. For a bias of 0.6, a sample of 50 uncovered discrimination less than 15 percent of the time, but the power increases sharply with sample size. For a sample of 200, the power is nearly 67 percent; for a sample of 400, the power is 94 percent. For our highest level of bias, our smallest sample size found discrimination 45 percent of the time, and samples over 250 uncovered discrimination every time.

Employing only the variables available in the HMDA data significantly lowered the chance of passing the examination (see bottom of table 5 and figure 2). One criticism of a test using this incomplete data is that it suffers from omitted variable bias. In our model, differences in credit history and assets result in a higher rejection rate for minorities, but the regression results attribute this to race even in the absence of discrimination. While this tends to make it easier to find bias when it exists, it also makes it easier to find bias when it does not exist. In the case of no bias, the power of the test plummets as the sample size increases, so that with sample sizes of 800, nondiscriminatory banks would pass less than a third of the time.

Given that one perceived advantage of using the available HMDA data is that large samples can be put together at low cost, this result suggests that indications of discrimination that rely solely on HMDA data should be treated with caution. The usefulness of employing the readily available HMDA data to pretest banks in order to direct scarce regulatory resources more effectively is a possible extension of our analysis.

Of course, the problem of false positives arises partly because of the built-in correlations between race and the other variables. If race and these variables were uncorrelated, the problem of false positives would be reduced. The degree of correlation between the two is an empirical question.

Initially, we ran ordinary least squares (OLS) instead of the more sophisticated logit model to save time in our simulations. While crude because it fails to adjust for the 0/1 nature of the dependent variable, the OLS test performs nearly as well as the logit estimator (see table 6). It also never fails to achieve convergence as the logit sometimes does with small sample sizes and a high degree of bias. Consequently, the OLS test using the HMDA data may be useful as a pretest when logit fails.

TABLE 6

**Proportion Passing Examination
(1,000 repetitions, DLS test)**

Sample size	Discrimination Parameter—Full Set of Variables ^a				
	0.0	0.2	0.4	0.6	0.8
50	0.974	0.942	0.923	0.806	0.358
75	0.968	0.953	0.886	0.694	0.173
100	0.970	0.928	0.851	0.635	0.089
150	0.969	0.922	0.832	0.453	0.014
200	0.959	0.900	0.776	0.361	0.005
250	0.957	0.895	0.736	0.225	0.000
300	0.942	0.887	0.684	0.156	0.000
350	0.951	0.863	0.634	0.111	0.000
400	0.932	0.859	0.562	0.097	0.000
450	0.935	0.834	0.528	0.055	0.000
500	0.924	0.825	0.490	0.038	0.000
600	0.909	0.785	0.415	0.021	0.000
800	0.898	0.748	0.306	0.003	0.000

Sample size	Discrimination Parameter—HMDA Variables ^a				
	0.0	0.2	0.4	0.6	0.8
50	0.904	0.864	0.791	0.657	0.245
75	0.873	0.829	0.710	0.479	0.117
100	0.860	0.782	0.664	0.399	0.039
150	0.819	0.720	0.540	0.208	0.008
200	0.758	0.674	0.425	0.131	0.000
250	0.731	0.567	0.336	0.043	0.000
300	0.650	0.512	0.269	0.030	0.000
350	0.641	0.443	0.191	0.012	0.000
400	0.582	0.384	0.140	0.006	0.000
450	0.568	0.336	0.113	0.002	0.000
500	0.519	0.301	0.072	0.000	0.000
600	0.430	0.226	0.046	0.000	0.000
800	0.317	0.112	0.017	0.000	0.000

a. If zero, no bias. If one, no loans made to minorities.
SOURCE: Authors' calculations.

IV. Analysis of Alternative Exam Procedures

A goal of this research is to test the power of actual examination techniques used in consumer exam settings. The *Federal Reserve System Consumer Compliance Handbook*, published by the Federal Reserve Board of Governors (1989), provides guidance on how to model a consumer exam. In addition, we have met with consumer examiners and one of us went on an actual consumer examination to observe procedures firsthand.

Applicant profile worksheets are the main tool used by consumer examiners to test for discrimination against protected classes such as minorities or females. Examiners complete these forms from a sample of both accepted and rejected loan files. They list the applicant's class characteristics along with his or her length of employment, length of residence, and monthly debt/income ratio. The forms also include the date and terms of the requested credit. If the application is rejected, reasons for rejection are noted.

The examiner then uses the profiles to compare the characteristics of applicants who receive credit with those who do not, and to make comparisons between protected classes. As a first check, the examiner sees whether those who are accepted or rejected are treated in accordance with the bank's articulated lending criteria. Any instances of credit decisions that fall outside the criteria are flagged for further investigation. The examiner then has considerable flexibility in how the files are selected and segregated for analysis between protected classes. Various comparisons suggested by the handbook include accepted minority versus accepted nonminority, rejected minority versus accepted nonminority, and rejected minority versus rejected nonminority. While not conducting a formal statistical test, the examiner then makes a judgment as to whether the classes have received equal treatment. With respect to sample size, the *Consumer Compliance Handbook* notes:

Since statistical validity is not a key issue, the ideal size of the judgmental sample cannot be stated in terms of numbers. Enough items should be selected in order to draw a reasonable conclusion. Again, the examiner should exercise careful discretion based upon experience and related examination findings. (p. 1.B.25)

Discussion with experienced examiners suggests that the examiner starts off with a fairly small sample size of perhaps 40 acceptances and 40 rejections. This small sample size is due in part to the limited amount of time available to conduct the examination. In addition, for many other of the regulations tested on an examination — such as truth in lending — compliance can be adequately ascertained through a small sample. If the examiner finds any evidence of discrimination — for example, a rejected minority whose characteristics dominated those of an accepted white — then the sample size is expanded and a more intensive investigation is conducted.

TABLE 7

**Proportion Passing Examination
(1,000 repetitions, NP I)**

Sample size	Discrimination Parameter—Full Set of Variables ^a				
	0.0	0.2	0.4	0.6	0.8
50	0.484	0.466	0.390	0.288	0.058
75	0.653	0.578	0.495	0.328	0.043
100	0.665	0.561	0.485	0.316	0.020
150	0.658	0.588	0.458	0.235	0.006
200	0.656	0.561	0.421	0.200	0.002
250	0.617	0.548	0.406	0.154	0.001
300	0.585	0.477	0.377	0.135	0.000
350	0.591	0.480	0.335	0.098	0.000
400	0.558	0.457	0.319	0.084	0.000
450	0.530	0.406	0.260	0.077	0.000
500	0.495	0.372	0.227	0.060	0.000
600	0.426	0.354	0.196	0.031	0.000
800	0.318	0.227	0.125	0.020	0.000

Sample size	Discrimination Parameter—HMDA Variables ^a				
	0.0	0.2	0.4	0.6	0.8
50	0.044	0.058	0.044	0.033	0.015
75	0.027	0.021	0.016	0.011	0.005
100	0.009	0.010	0.008	0.001	0.001
150	0.001	0.002	0.002	0.000	0.000
200	0.001	0.000	0.002	0.000	0.000
250	0.000	0.000	0.000	0.000	0.000
300	0.000	0.000	0.000	0.000	0.000
350	0.000	0.000	0.000	0.000	0.000
400	0.000	0.000	0.000	0.000	0.000
450	0.000	0.000	0.000	0.000	0.000
500	0.000	0.000	0.000	0.000	0.000
600	0.000	0.000	0.000	0.000	0.000
800	0.000	0.000	0.000	0.000	0.000

a. If zero, no bias. If one, no loans made to minorities.
SOURCE: Authors' calculations.

As an approximation of actual exam procedures, we tested the power of two potential exam procedures on our generated loan files. Our first test (NP 0) looked for *any* instance of a rejected minority with financial characteristics that dominated those of an accepted white. Domination was defined as having more favorable characteristics for all four of the criteria used in the loan scoring procedures. The second test (NP I) compared the proportion of minority rejections that dominated white acceptances with the proportion of minority acceptances that dominated white rejections. A third test attempted to see whether differences in the latter proportions are statistically significant.

NP 0 proved to create a high degree of false positives. With BIAS set equal to zero and a sample size of 50, the exam indicated discrimination 39 percent of the time (with 100 repetitions). For a sample size of 100, discrimination was found 72 percent of the time. Finally, for sample sizes above 200, we almost always found an instance of a minority rejected applicant who had more favorable financial characteristics than an accepted white. This high degree of false positives suggests that the test was overly stringent given the degree of randomness we introduced in the loan files. Whenever there is even a small amount of randomness in the approval process, the probability of finding a rejected minority applicant who dominated an approved white applicant approaches one, so this test actually performs worse as the sample size expands in the case where there is no bias. On the other hand, if there is no randomness and all of the variables are measured without error, then this test would perform flawlessly. These are conditions that are unlikely to be met with actual data.

We used NP I to account for the underlying uncertainty in the credit approval process. We compared the proportion of minority rejections that were dominated by white acceptances with the proportion of white rejections that were dominated by minority acceptances. If the first proportion was larger, we took this as a flag for discrimination. The power of this exam is shown in table 7 for sample sizes of 50 through 800.

For BIAS set equal to 0, we find that this test reports a large proportion of false positives, better than NP 0, but much worse than the logit and OLS tests. Using only the HMDA variables resulted in false positives almost all of the time regardless of the sample size. When bias is introduced, this test outperforms logit and OLS in small samples, but not in large samples. Unfortunately, the large proportion of false positives in the case of no bias makes this test less than ideal.

In table 8, we report results for a modified version of this test (NP II) that attempts to determine whether differences between the proportion of minority rejections that dominated some whites and white rejections that were dominated by some minorities was statistically significant using a chi-squared test. When there is no bias, this test has fewer false positives in small sample sizes than NP I, but has more when the sample size is large. Like NP I, when there is bias, the test is much better at detecting it than are logit and OLS in small samples, but the test is not as good with large samples.

TABLE 8

Proportion Passing Examination
(1,000 repetitions, NP II)

Sample size	Discrimination Parameter—Full Set of Variables ^a				
	0.0	0.2	0.4	0.6	0.8
50	0.703	0.684	0.685	0.640	0.538
75	0.629	0.626	0.570	0.550	0.403
100	0.545	0.537	0.479	0.458	0.280
150	0.420	0.406	0.395	0.346	0.139
200	0.405	0.344	0.361	0.309	0.078
250	0.339	0.331	0.329	0.252	0.042
300	0.337	0.322	0.275	0.256	0.025
350	0.262	0.283	0.289	0.236	0.022
400	0.285	0.270	0.260	0.197	0.017
450	0.279	0.275	0.249	0.190	0.004
500	0.237	0.227	0.224	0.184	0.005
600	0.211	0.210	0.198	0.163	0.005
800	0.175	0.184	0.179	0.164	0.001

Sample size	Discrimination Parameter—HMDA Variables ^a				
	0.0	0.2	0.4	0.6	0.8
50	0.749	0.798	0.753	0.728	0.742
75	0.741	0.727	0.707	0.689	0.640
100	0.671	0.651	0.679	0.636	0.612
150	0.584	0.615	0.608	0.539	0.509
200	0.555	0.554	0.556	0.503	0.490
250	0.511	0.538	0.532	0.502	0.385
300	0.480	0.493	0.469	0.470	0.324
350	0.466	0.452	0.471	0.465	0.331
400	0.420	0.432	0.450	0.446	0.314
450	0.422	0.422	0.441	0.449	0.284
500	0.400	0.409	0.399	0.413	0.282
600	0.350	0.366	0.376	0.402	0.239
800	0.318	0.330	0.338	0.382	0.212

a. If zero, no bias. If one, no loans made to minorities.
SOURCE: Authors' calculations.

V. Conclusion

Using a simulation model, we have examined several approaches to testing whether a financial institution discriminates. Because we employ a simulation model, the degree of bias can be varied from no bias to the point where no minorities are given loans.

Tests that employ all of the information included in our simulated loan files perform much better than those that use only the HMDA subset of data. For example, using the logit test, a nondiscriminating bank with 800 applications has less chance of passing than a smaller discriminating bank (bias = 0.4) with only 250 applications (see table 5). More surprisingly, low levels of bias can be difficult to detect even with large sample sizes. With levels of apparent bias found in actual HMDA data, sample size is very important. Tests with sample sizes under 50 almost always fail to detect bias, whereas tests with sample sizes greater than 200 perform well. Our test that attempts to mimic the procedures employed by examiners suggests that they work well in small samples, but also tend to find bias even in simulations when it is not present.

The qualitative characteristics of these findings are unlikely to be affected by either better calibration of the data or more elaborate modeling of the approval process. Detecting bias, particularly a small degree of bias at an institution, is likely to be a difficult endeavor. Even examiners, who have access to the applicants' loan files, are apt to face problems. Statistical methods require large sample sizes for low bias levels, which may require a great deal of regulatory resources. Examiner-inspired methods work well in small samples, but have a tendency to find bias even when it is not present. In particular, any randomness in lending decisions makes simple match-pair tests (such as NP 0) yield a high degree of false positives. More sophisticated versions (such as NP I and NP II) perform better because they allow for some underlying randomness.

Future research will look at the usefulness of employing the HMDA variables as a pretest to direct regulatory resources. By construction, this paper cannot say whether there is discrimination in mortgage lending, but by laying out the issues and problems involved in testing for discrimination and by exploring the robustness of the various approaches to testing for bias, it allows a more informed debate to proceed.

References

- Avery, Robert B., Patricia E. Beeson, and Mark S. Sniderman. "Lender Consistency in Housing Credit Markets," Federal Reserve Bank of Cleveland. Working Paper 9309, December 1993.
- Board of Governors of the Federal Reserve System, Division of Consumer and Community Affairs. *Federal Reserve System Consumer Compliance Handbook*, Washington, D.C., June 1989.
- Federal National Mortgage Association, *Fannie Mae Guides*, vol. 1 (Selling Guide), Washington, D.C., 1992.
- Horne, David K. "Evaluating the Role of Race in Mortgage Lending," *FDIC Banking Review*, vol. 7, no. 1 (Spring/Summer 1994), pp. 1-15.
- Kennickell, Arthur, and Janice Shack-Marquez. "Changes in Family Finances from 1983 to 1989: Evidence from the Survey of Consumer Finances," *Federal Reserve Bulletin*, vol. 78, no. 1 (January 1992), pp. 1-18.
- Munnell, Alicia H., Lynn E. Browne, James McEaney, and Geoffrey M.B. Tootell. "Mortgage Lending in Boston: Interpreting HMDA Data," Federal Reserve Bank of Boston, Working Paper No. 92-7, October 1992.

Third Quarter Working Papers

Current *Working Papers* of the Cleveland Federal Reserve Bank are listed in each quarterly issue of the *Economic Review*. Copies of specific papers may be requested by completing and mailing the attached form below.

Single copies of individual papers will be sent free of charge to those who request them. A mailing list service for personal subscribers, however, is not available.

Institutional subscribers, such as libraries and other organizations, will be placed on a mailing list upon request and will automatically receive *Working Papers* as they are published.

■ **9408**
**The Impact of the
AFDC Program on Birth
Decisions and Program
Participation**
by Elizabeth T. Powers

■ **9409**
**The Welfare Effects
of Tax Simplification:
A General-Equilibrium
Analysis**
by Jang-Ting Guo and
Kevin J. Lansing

■ **9410**
**Tax Structure, Welfare,
and the Stability of
Equilibrium in a Model
of Dynamic Optimal
Fiscal Policy**
by Jang-Ting Guo and
Kevin J. Lansing

Please complete and detach the form below and mail to:
Research Department
Federal Reserve Bank of Cleveland
P.O. Box 6387
Cleveland, Ohio 44101

Check item(s) requested

Send to:
Please print

Please send the following Working Paper(s):

9408 **9409** **9410**

Name

Address

City

State

Zip

Working Papers

- **9301**
Sharing with a Risk-Neutral Agent
 by Joseph G. Haubrich
- **9302**
HRM Policy and Increasing Inequality in a Salary Survey
 by Erica L. Groshen
- **9303**
Regulatory Taxes, Investment, and Financing Decisions for Insured Banks
 by Anlong Li, Peter Ritchken, L. Sankarasubramanian, and James B. Thomson
- **9304**
Measuring Core Inflation
 by Michael F. Bryan and Stephen G. Cecchetti
- **9305**
Generational Accounting in Norway: Is the Nation Overconsuming Its Petroleum Wealth?
 by Alan J. Auerbach, Jagadeesh Gokhale, Laurence J. Kotlikoff, and Erling Steigum, Jr.
- **9306**
The Evolving Legal Framework for Financial Services
 by Walker F. Todd
- **9307**
Loan Sales, Implicit Contracts, and Bank Structure
 by Joseph G. Haubrich and James B. Thomson
- **9308**
Dynamic Optimal Fiscal and Monetary Policy in a Business Cycle Model with Income Redistribution
 by Kevin J. Lansing
- **9309**
Lender Consistency in Housing Credit Markets
 by Robert B. Avery, Patricia E. Beeson, and Mark S. Sniderman
- **9310**
Accounting for Racial Differences in Housing Credit Markets
 by Robert B. Avery, Patricia E. Beeson, and Mark S. Sniderman
- **9311**
The Equity of Social Services Provided to Children and Senior Citizens
 by Laurence J. Kotlikoff and Jagadeesh Gokhale
- **9312**
Business Cycles and Aggregate Labor-Market Fluctuations
 by Finn E. Kydland
- **9313**
Loan Sales as a Response to Market-Based Capital Constraints
 by Charles T. Carlstrom and Katherine A. Samolyk
- **9314**
Accounting for Earnings Inequality in a Diverse Workforce
 by Mark E. Schweitzer
- **9401**
Exclusion in All-Pay Auctions
 by Ian Gale and Mark Stegeman
- **9402**
Auctions with Budget-Constrained Buyers: A Nonequivalence Result
 by Yeon-Koo Che and Ian Gale
- **9403**
Underlying Determinants of Closed-Bank Resolution Costs
 by William P. Osterberg and James B. Thomson
- **9404**
Depositor Preference and the Cost of Capital for Insured Depository Institutions
 by William P. Osterberg and James B. Thomson
- **9405**
The Federal Reserve Board before Marriner Eccles
 by Walker F. Todd
- **9406**
Optimal Fiscal Policy when Public Capital Is Productive: A Business Cycle Perspective
 by Kevin J. Lansing
- **9407**
Anticipating Bailouts: The Incentive-Conflict Model and the Collapse of the Ohio Deposit Guarantee Fund
 by Ramon P. DeGennaro and James B. Thomson
- **9408**
The Impact of the AFDC Program on Birth Decisions and Program Participation
 by Elizabeth T. Powers
- **9409**
The Welfare Effects of Tax Simplification: A General-Equilibrium Analysis
 by Jang-Ting Guo and Kevin J. Lansing
- **9410**
Tax Structure, Welfare, and the Stability of Equilibrium in a Model of Dynamic Optimal Fiscal Policy
 by Jang-Ting Guo and Kevin J. Lansing

Economic Review

■ 1993 Quarter 3

**Capital Requirements and Shifts
in Commercial Bank Portfolios**
by Joseph G. Haubrich and
Paul Wachtel

FDICIA's Emergency Liquidity Provisions
by Walker F. Todd

**Efficiency and Technical Progress
in Check Processing**
by Paul W. Bauer

■ 1993 Quarter 4

Required Clearing Balances
by E.J. Stevens

The CPI as a Measure of Inflation
by Michael F. Bryan and
Stephen G. Cecchetti

**The Inaccuracy of Newspaper Reports
of U.S. Foreign Exchange Intervention**
by William P. Osterberg and
Rebecca Wetmore Humes

■ 1994 Quarter 1

Institutional Aspects of U.S. Intervention
by Owen F. Humpage

**The 1995 Budget and Health Care Reform:
A Generational Perspective**
by Alan J. Auerbach, Jagadeesh Gokhale, and
Laurence J. Kotlikoff

**On Disinflation since 1982:
An Application of Change-Point Tests**
by Edward Bryden and John B. Carlson

■ 1994 Quarter 2

**U.S. Banking Sector Trends: Assessing
Disparities in Industry Performance**
by Katherine A. Samolyk

**Competition for Scarce Inputs: The Case of
Airport Takeoff and Landing Slots**
by Ian Gale

**Regional Wage Convergence and Divergence:
Adjusting Wages for Cost-of-Living Differences**
by Randall W. Eberts and Mark E. Schweitzer

Economic Commentary

Free Markets and Price Stability: Opportunities for Mexico

by Jerry L. Jordan
August 1, 1993

Assessing Real Interest Rates

by John B. Carlson
August 15, 1993

Airline Deregulation: Is It Time to Finish the Job?

by Paul W. Bauer and Ian Gale
September 1, 1993

The Decline in U.S. Saving Rates: A Cause for Concern?

by Jagadeesh Gokhale
September 15, 1993

Credibility Begins with a Clear Commitment to Price Stability

by Jerry L. Jordan
October 1, 1993

The Budget Reconciliation Act of 1993: A Summary Report

by David Altig and
Jagadeesh Gokhale
October 15, 1993

Making the SAIF Safe for Taxpayers

by William P. Osterberg and
James B. Thomson
November 1, 1993

Community Lending and Economic Development

by Jerry L. Jordan
November 15, 1993

Replacing Reserve Requirements

by E.J. Stevens
December 1, 1993

Long-Term Health Care: Is Social Insurance Desirable?

by Jagadeesh Gokhale and
Lydia K. Leovic
December 15, 1993

Monetary Policy and Inflation: 1993 in Perspective

by Gregory A. Bauer and
John B. Carlson
January 1, 1994

Report of the Fourth District Economists' Roundtable

by Michael F. Bryan and
John B. Martin
January 15, 1994

Are Service-Sector Jobs Inferior?

by Max Dupuy and
Mark E. Schweitzer
February 1, 1994

The National Depositor Preference Law

by James B. Thomson
February 15, 1994

Issues in CRA Reform

by Mark S. Sniderman
March 1, 1994

Back to the Future: A View of Prospective Deficits through the Prism of the Past

by David Altig and
Jagadeesh Gokhale
March 15, 1994

Central Bank Independence

by Owen F. Humpage
April 1, 1994

Health Care Reform from a Generational Perspective

by David Altig and
Jagadeesh Gokhale
April 15, 1994

Lessons from the Collapse of Three State-Chartered Private Deposit Insurance Funds

by Walker F. Todd
May 1, 1994

Assessing Progress toward Price Stability: Looking Forward and Looking Backward

by John B. Carlson
May 15, 1994

The Government's Role in the Health Care Industry:

Past, Present, and Future
by Charles T. Carlstrom
June 1, 1994

Are the Japanese to Blame for Our Trade Deficit?

by Owen F. Humpage
June 15, 1994

A Beginner's Guide to the U.S. Payments System

by Paul W. Bauer
July 1, 1994

Must the Fed Fight Growth?

by Jerry L. Jordan
July 15, 1994

Midyear Report of the Fourth Dis- trict Economists' Roundtable

by Michael F. Bryan and
John B. Martin
August 1, 1994

Looking Back at Slow Employment Growth

by Kristin M. Roberts and
Mark E. Schweitzer
August 15, 1994

Banking and the Flow of Funds: Are Banks Losing Market Share?

by Katherine A. Samolyk
September 1, 1994

The Economics of Health Care Reform

by Charles T. Carlstrom
September 15, 1994

Federal Reserve
Bank of Cleveland
Research
Department
P.O. Box 6387
Cleveland, OH 44101

BULK RATE
U.S. Postage Paid
Cleveland, OH
Permit No. 385

Address Correction

Requested:

Please send

corrected

mailing label to the

Federal Reserve

Bank of Cleveland

Research

Department

P.O. Box 6387

Cleveland OH 44101