

A Review of Statistical Problems in the Measurement of Mortgage Market Discrimination and Credit Risk

Anthony M. Yezer



A Review of Statistical Problems in the Measurement of Mortgage Market Discrimination and Credit Risk

Anthony M. Yezer

September 2010

Research Institute for Housing America

BOARD OF TRUSTEES

Chair

Teresa Bryce, Esq.
Radian Group Inc.

Michael W. Young
Cenlar FSB

Nancee Mueller
Wells Fargo

Edward L. Hurley
Avanath Capital Partners LLC

Steve Graves
Principal Real Estate Investors

Dena Yocom
IMortgage

STAFF

Jay Brinkmann, Ph.D.
Senior Vice President, Research and Business Development
Chief Economist
Mortgage Bankers Association

Michael Fratantoni, Ph.D.
Vice President, Research and Economics
Mortgage Bankers Association

Table of Contents

| | |
|---|-----------|
| Executive Summary | 7 |
| I. Introduction | 9 |
| II. General Comments: Taking the “Con” out of Economics | 13 |
| III. Conventional Approaches to Measurement of Mortgage Market Discrimination and Credit Risk | 17 |
| A. Conventional Approaches to Testing for Discrimination | |
| 1. Testing for Discrimination Based on Applicant Rejection Equations | |
| 2. Testing for Discrimination Based on Mortgage Pricing Equations | |
| 3. Testing for Discrimination Based on Mortgage Default Equations | |
| B. Conventional Approaches to Measurement of Credit Risk | |
| IV. Problems with Conventional Approaches to Measurement of Mortgage Market Discrimination and Credit Risk | 29 |
| A. Sources of Bias in Conventional Statistical Approaches | |
| 1. Omitted Variable Bias | |
| 2. Other Problems | |
| B. Lack of Theoretical Support for Conventional Statistical Approaches | |
| C. Theoretical Support for Valid Statistical Models of Discrimination and Credit Risk in Mortgage Lending | |
| D. Critique of Conventional Models of Mortgage Market Discrimination | |
| E. Critique of Conventional Models of Mortgage Credit and Payment Risk | |
| V. Conclusions | 49 |
| End Notes | 51 |

Executive Summary

Over the past twenty years, understanding of and business practice in mortgage markets has been influenced significantly by the application of statistical models. Mortgage underwriting was automated using statistical models of default and default loss, and statistical models of denial rates and loan pricing were used to test for discrimination in lending. Efforts to measure mortgage market discrimination and credit risk have been propelled by an increase in the loan-level data available through various resources. Unfortunately, as researchers strived to produce results from these data, critical statistical errors were overlooked and then repeated in what has become the “conventional approach” to measuring discrimination and credit risk. The purpose of this paper is to re-examine the fundamental assumptions integrated into this conventional model and provide insight into why the results are both biased and inaccurate.

This study will argue that conventional statistical models of discrimination and mortgage credit lack a sound basis in economic theory and rely on unrealistic and demonstrably false assumptions. As a result of these shortcomings, discrimination tests tend to produce false-positive indications of discrimination where none exists, and tests for default risk fail to predict instances where default rates are likely to rise significantly.

A common theme underlies this essay: the mortgage lending transaction is extremely complex and involves many dimensions. Applicants, loan officers, underwriters and secondary market participants make decisions based on simultaneous consideration of many factors about which both the applicant and the lender must come to some mutual agreement. Applicants choose among mortgage lenders, products and terms based on their personal circumstances, with higher risk applicants self-selecting into loan programs with higher mortgage rates and higher rejection and default rates. These higher rejection and default rates are due to their self-selection into particular loan programs, not to differential treatment by lenders. The problem with conventional statistical techniques for estimating mortgage discrimination and credit risk is that these methods assume that borrowers never consider the effects of their decisions on the mortgage transaction. You do not need to be an economist to understand that mortgage applicants behave strategically when choosing mortgage products.

These critiques of conventional models of mortgage market discrimination and credit risk are not new. The difficulty with such findings is that they also imply that there is no easy way to test for discrimination in mortgage markets or to estimate credit and prepayment risk. Given the demand for testing and estimation related to discrimination and credit risk, current models have been sold as a low-budget answer to a difficult problem. Unfortunately this “solution” often gives unreliable and profoundly misleading results. (Recent experience clearly contradicts the notion that performance of mortgages is well described by simple default and prepayment models. Put another way, the experience of housing and mortgage markets in the United States since 2006 leads to the conclusion that something seems to have gone terribly wrong.)

The serious limitations of current statistical approaches to testing for discrimination and credit risk in mortgage lending have likely contributed to recent problems in mortgage markets. If these limitations are not recognized and naïve reliance on them continues, current problems are likely to recur in the future. Alternatively, there are major gains to be made if economic analysis of mortgage market discrimination and mortgage credit risk can be improved.

I. Introduction

Statistical and econometric analysis currently plays a major role in the measurement of mortgage market discrimination and credit risk. A major revolution in consumer finance was based on the use of statistical techniques to construct credit scoring schemes that are now applied to mortgage lending. These quantitative techniques can be confusing and intimidating for many people. Indeed, many professional economists, managers and lawyers frequently have difficulty understanding the methods and implications of statistical analysis. In economics, debates often revolve around different views regarding appropriate statistical methods, as opposed to philosophical or ideological differences. Unfortunately, statistical errors that are covered in undergraduate courses are sometimes made by faculty submitting papers for publication, consultants preparing an analysis for trial, regulators examining lenders and business analysts developing a strategy. These errors are not always obvious or careless, but the effects of even small errors can be profound.

Statistical errors that are not initially detected and corrected can easily be perpetuated with one methodological approach being replicated once it becomes “conventional practice.” Indeed, the fundamental rationale for and weaknesses of a conventional approach are often forgotten once it has become popularized.

This has become the case in what will be termed here “conventional” approaches to testing for mortgage market discrimination and measuring mortgage credit risk. The purpose of this paper is to step back and re-examine the fundamental assumptions regarding how statistical analysis of discrimination and credit risk is performed.

In spite of overwhelming evidence of their flaws, conventional approaches to measuring discrimination and credit risk in mortgage lending are still used today. What are the characteristics of the conventional approaches? They are simple statistical models usually involving a single equation that are not firmly grounded in economic theory. The major simplifying assumption made in these models is that borrowers have no knowledge of the mortgage lending process and do not select mortgage terms strategically.

These assumptions are inconsistent with models of the mortgage lending process that have been in the academic economics literature since the early 1980's. As a general rule in economics, whenever empirical work ignores or is inconsistent with economic theory, major problems arise in the interpretation of any empirical results that are obtained. Conventional models of mortgage lending are an excellent example of this general principle.

How did these conventional empirical approaches gain status with bankers, regulators, and the public in general? First they followed upon significant innovations in the use of statistical techniques to measure various aspects of mortgage lending, including tests for discrimination and credit risk. The advance of these statistical models was facilitated by automation of the mortgage lending and servicing processes which has made detailed data on both mortgage lending decisions and subsequent mortgage performance available for analysis. As is the case when massive datasets are applied to any problem of statistical inference, very precise estimates of model parameters can be obtained. This has led some to the conclusion that the determinants of lender behavior in making credit decisions and behavior of mortgages once endorsed are well described by simple single equation models.

However, recent experience appears to contradict the notion that performance of mortgages is well described by simple default and prepayment models and losses, particularly on alt-A and subprime lending. Put another way, the experience of housing and mortgage markets in the United States since 2006 leads to the conclusion that something seems to have gone terribly wrong.

While conventional models of mortgage discrimination, prepayment, and default achieve high levels of apparent precision because they use large quantities of high quality data, they are notably lacking in theoretical support. The major point of this paper is that theoretical models of mortgage lending, default and prepayment processes imply that the statistical techniques used to estimate these models of discrimination and credit risk produce biased and inconsistent results.

The theoretical problem with conventional statistical techniques for estimating mortgage discrimination and credit risk arises because these methods assume that borrowers never consider the effects of their decisions on the mortgage transaction. Put another way, conventional statistical techniques assume that borrowers never behave strategically and mortgage terms are selected by the lender. This contrasts with standard economic theory which suggests that rational borrowers consider the effects of decisions regarding home value, mortgage amount, monthly payment, down payment, cosigners, prepayment penalties, etc. on any aspect of the mortgage transaction including likelihood of rejection, interest rate, points, APR, prepayment or default. Conversely, conventional statistical techniques for measuring discrimination, prepayment and default assume that borrowers select mortgage terms without regard to any outcome of the mortgage application process and the loan pricing decision.

For example, conventional statistical techniques assume that borrowers determine the amount that they are willing to pay for a home without considering the effect of this decision on the chances of rejection, the cost of credit and future prepayment or default. To anyone who has discussed a home

purchase with a realtor, this assumption may seem fantastic because, among the first concerns of a competent realtor is the ability of the homebuyer to qualify for a mortgage.¹ In the case of refinancing, the statistical models require the equally unlikely assumption that the owner does not consider either likelihood of rejection, future possibility of default or effects on the APR when determining the amount of any cash out realized in connection with the refinancing. Clearly, these conventional statistical models of mortgage discrimination and credit risk rely on assumptions that are not only at variance with economic theory of rational consumer choice, they also disagree with common perceptions of borrower behavior and typical borrower experience. Put another way, you do not need to be an economist to know that mortgage applicants behave strategically when choosing mortgage products.

The fact that conventional statistical models of discrimination and mortgage credit lack a sound basis in economic theory and rely on unrealistic and demonstrably false assumptions is not just an academic issue. This essay will demonstrate that these assumptions introduce systematic biases into the estimates that make the models fail in ways that are particularly troubling. Discrimination tests tend to produce false positive indications of discrimination when none exists and tests for default risk are particularly bad at detecting instances when future default rates are likely to rise significantly.

The remainder of this paper is divided into four sections. In the next section, general comments regarding problems in empirical research in economics are discussed briefly. The third section reviews conventional statistical approaches to measurement of mortgage market discrimination and credit risk. Several different methods are used to measure discrimination, including mortgage rejection, pricing and default equations. Somewhat different statistical models are used to measure the probability of default and prepayment. This section shows that these methods, while appearing different on the surface, share common assumptions about the mortgage underwriting process. The fourth section discusses serious flaws that are common to all of these conventional statistical models of discrimination and mortgage credit risk. Curiously, the same flaws produce false indications of discrimination in some models and of non-discrimination in others. Models of credit risk are seriously compromised precisely during periods when credit risk is highest, such as the period just experienced in mortgage markets. The final section presents conclusions and some recommendations for the measurement of discrimination and credit risk in mortgage markets.

II. General Comments: “Taking the Con out of Economics”

There is a long-standing debate in economics over the proper use of data and statistical techniques particularly where the results have important policy or practical implications. Perhaps the most famous paper in this literature is Leamer (1983), “Let’s Take the Con Out of Econometrics” in which he observed that: “Hardly anyone takes data analysis seriously. Or, perhaps more accurately, hardly anyone takes anyone else’s data analysis seriously.”² In fact, complaints about statistical testing without a firm basis in theory that are at the heart of this essay on mortgage markets, extend back to Haavelmo (1944) who lamented:

“A design of experiments (a prescription of what the physicists call a ‘crucial experiment’) is an essential appendix to any quantitative theory. And we usually have some such experiment in mind when we construct the theories, although – unfortunately – most economists do not describe their design of experiments explicitly.” (pg. 14)

This quotation is directly relevant to models of discrimination and credit risk because Haavelmo received the Nobel Prize in Economics primarily for developing statistical estimators that could be used to estimate multiple-equation models and the major critique of mortgage models made in this essay is that, 65 years after Haavelmo, they continue to rely on single-equation models and ignore the problem of simultaneous equation bias that prompted his research.

Recently the *Journal of Economic Perspectives* published papers by prominent econometricians from a symposium on the topic “Con Out of Economics.” The lead paper by Angrist and Pischke (2010) notes that economists are still struggling with the problem of making empirical inferences from non-experimental data. They see progress in some areas where better research designs provide the basis for statistical inference and find fault with current approaches in other areas. In the same volume, Leamer (2010) is particularly critical of the application of econometric models of credit and prepayment risk and their use in mortgage underwriting and pricing securities.

Much of the discussion in the symposium uses the literature on the deterrent effects of capital punishment as an example. Early papers in this literature used single equation models in which the murder rate in a state was related to the existence of capital punishment or the number of executions. These simple models produced either a non-significant effect of capital punishment or a positive coefficient which, if taken literally implies that the deterrent effect is negative. Ehrlich (1975) noted that single equation models were flawed and argued that, while capital punishment influenced the murder rate, it was also true that higher murder rates led to the adoption of capital punishment and increased the frequency of its use. Thus higher murder rates could also cause increased reliance on capital punishment and this reverse causality could be confounding the statistical analysis.³ This is a classic example of a situation where the estimated coefficient of a capital punishment variable in a murder rate equation is biased upward, i.e. tends to be positive when the true effect of punishment on crime is negative, due to simultaneous equations bias. Ehrlich pointed out the problem, and, based on economic theory, specified a three-equation model of murder, capital punishment and enforcement. His estimates indicated that, in contrast to the single equation models, the effect of capital punishment on murder was negative and statistically significant. These results were influential in debates regarding public policy toward the death penalty.

Subsequent literature on the relation between the death penalty and murder rates argued that Ehrlich was correct about murder, punishment and enforcement being jointly determined but that there was little theoretical support for the particular three-equation system that he estimated. These subsequent papers found that changes in the variables included in the model could change estimates of the effects of capital punishment on the murder rate dramatically.

What is the current status of the debate over the effects of capital punishment on murder? Based on the discussion in the symposium articles, it appears that murder, death penalty and enforcement are jointly determined variables and single equation models produce biased estimates of the effects of the death penalty. Furthermore, given the lack of a firm theoretical basis for identifying variables that can be used to estimate a multi-equation system and the poor quality of the data, firm estimates of the effects of the death penalty on murder rates are not currently available.

This may seem to be a very inadequate and disappointing outcome given the amount of research effort on the death penalty and the importance of the debate for public policy. However, as this essay will demonstrate, the current state of statistical testing for mortgage discrimination and credit risk is even less advanced than the literature on effects of capital punishment. Much of the literature on mortgage lending still uses single equation models of the mortgage lending process which have no firm basis in economic theory. Put another way, much of the current literature ignores the insights from Ehrlich (1975). Accordingly, we are a long way from getting the con out of economic analysis of mortgage discrimination and credit risk. The succeeding sections of this essay will provide a detailed demonstration of problems with current methods and suggest the direction necessary for their resolution.

III. Conventional Approaches to Measurement of Mortgage Market Discrimination and Credit Risk

This section is divided into two major subsections. The first discusses empirical techniques used in measurement of discrimination and the second discusses credit risk in the form of default and prepayment hazards.

III.A Conventional Approaches to Testing for Discrimination

There are three different versions of the conventional approach to testing for discrimination in mortgage lending. They involve statistical estimation of equations in which the dependent variable is either rejection of the mortgage application, mortgage pricing (APR) or mortgage default. In all three cases, the analysis involves estimation of a single equation model in which the dependent variable measuring loan outcome (rejection), loan pricing or loan default is regressed on a series of variables reflecting loan terms, financial characteristics of the applicant, characteristics of the real property collateral, and variables reflecting demographic factors, including minority status of the applicant or borrower.

III.A.1 Testing for Discrimination Based on Applicant Rejection Equations

The first example of a test for discrimination using a single equation model of rejection is Black, Schweitzer and Mandell (1978) who relied on a massive survey of banks and borrowers to get information on applicant rejection, loan terms, applicant financials and demographic characteristics. The authors find that two protected variables, applicant race and age, have a positive and significant relation to rejection for mortgages and home improvement loans. They provide no theoretical model of borrower behavior and caution that the effects are only significant at the ten percent confidence level. Curiously, the authors also warn that testing for discrimination using single equation models of interest rate or APR should not be attempted because terms of the loan are simultaneously determined.⁴ Thus this initial paper which served as the basis of the conventional approach to using rejection equations to test for discrimination has no theoretical basis and even concedes that loan terms such as APR and LTV are selected simultaneously by borrowers. Once it is conceded that borrowers select their down

payment in order to control their APR, the assumption that they do not use the down payment to control the probability of being rejected seems absurd. Borrowers should be at least as sensitive to rejection as they are to the APR offered conditional on acceptance.

In the 1980's the problem of discrimination in mortgage lending was dominated by concern with the fiscal viability of mortgage lenders. HMDA data had given researchers access to massive datasets documenting the rejection decisions of commercial banks but it was recognized that lack of information on both loan terms and borrower finances made estimates of the rejection produce false positive indications of discrimination due to omitted variables bias.⁵ Failure to include variables reflecting financial condition of the borrower, particularly credit score, would artificially cause any variable positively associated with credit score to appear to be negatively associated with the probability of rejection. For example, minority groups with credit scores below average should be rejected at a higher rate than non-minorities holding reported income, loan amount, residential location and other HMDA variable constant. The higher minority rejection rate, in this case, could be due either to discrimination or to the lower credit scores of the average minority borrower. This type of "spurious correlation" is the product of statistical analysis where important variables are omitted from the data.

The obvious inadequacy of HMDA data was well recognized and motivated the Federal Reserve Bank of Boston (Boston FED) to request cooperation of banks in the Boston MSA to aid in assembling a loan file-level dataset that would include details of mortgage terms and borrower characteristics. In theory, coding of all information in the loan file that was used by the underwriter would eliminate all omitted variable bias. In practice, it is difficult to reduce underwriting variables to a standardized dataset for two reasons. First, different lenders measure variables differently. Second, an important activity of the underwriter is verification of the key variables in the application. Failure to verify information in the application is a common reason for rejection but there is no standard for reporting such failure. As a practical matter, a rejection due to "unverifiable information" in the loan file can literally mean that the underwriter was unable to confirm the applicant's claims, it can indicate errors in the details of the claims, or it can mean that the underwriter had evidence that the claims were fraudulent. The difference between loans that are endorsed with unverified information and those rejected with the same notation is likely to be very large but omitted from the dataset.

Results of the Boston FED study appeared as an October 1992 working paper entitled "Mortgage Lending in Boston: Interpreting HMDA Data."⁶ The authors reference Black, Schweitzer and Mandell (1978) but make their own argument for estimating a rejection equation of the following form.⁷

$$P(D) = f(F, R, L, T, C) \quad (1)$$

where P(D) is the probability of denial, F is a vector of variables reflecting financial capacity of the applicant, R is variables measuring risks of default, L is measures of potential loss in default, T is a vector indicating loan terms and C is indicators of personal characteristics of the applicant, including

race, that are the object of the test for discrimination. The authors note that, using HMDA data alone, most variables would be eliminated from the model in equation (1) and the effect of race on rejection would be substantial. Adding variables, in their view, reduced omitted variable bias and lowered the differential in rejection rates between minority (African-American or Hispanic) and white applicants but statistically significant differences in rejection rates persisted.

Since it first appeared in 1992, the Boston FED study has had a number of critics and defenders. Just as the banks generously cooperated in providing the initial data, the Boston FED made it available to other researchers. Some have found that the results are not robust to changes in the specification of the equation while others find them convincing. It is useful to review some of the points made in the substantial list of papers that comprise this research but which have not dealt with the most fundamental problems of the method used in the Boston FED study.

Because the Boston FED study claimed to be solving the problem of omitted variable bias that characterized studies using HMDA data, a number of authors estimated alternative versions of equation (1) using different subsets of the many variables collected for the study. A series of papers, Carr and Megbolugbe (1993), Glennon and Stengel (1994) and Hunter and Walker (1996) showed that the positive and statistically significant dummy variable for minority applicant persisted when different groups of variables were swapped in and out of the model. In contrast, Harrison (1998) found that merely by forcing all of the variables collected by the Boston FED into the estimating equation, the estimated coefficient of the minority variable became non-significant.

Other researchers identified coding errors in the Boston FED data. Day and Liebowitz (1998) found that, correcting observations in the data based on logical consistency and dropping some cases where interest rates appear unreasonable, caused the estimated coefficient of the minority dummy variable to be non-significant. Horn (1997) reported the results from a major FDIC investigation that covered more than half of the lenders in the study. Trained examiners pulled individual loan files in an attempt to verify the data. They found numerous data coding errors including mischaracterization of the underwriting decision, factors considered by underwriters that were either omitted or incorrectly coded. There was also evidence of extra complexity in the underwriting process that was not modeled in the study. One significant source of errors was the difference between what was initially claimed by the applicant and the final determination of the underwriting process. Apparently, in cases where the applicant's claims of creditworthiness were contradicted by the findings of the underwriter, the false claims of the applicant were still recorded in the data. This is a common problem in analysis of mortgage application data because a major factor in loan denial is the inability to verify claims made by the applicant. Thus denied applications are more likely to have "unverified" information where unverified may be a euphemism for false claims by the applicant. In such cases, the information recorded in the loan file may include initial information provided by the applicant, corrected information from the underwriter or both. It is not clear what should be recorded in the data set used for statistical analysis when such contradictions occur.

Another problem concerns the very definition of a “loan rejection.” Horne (1997) reports that lenders often make counteroffers to applicants. If these counteroffers are accepted by the applicant, they are reported as acceptances and if the applicants do not take the counteroffers, they are reported as rejections. Clearly this is a case in which rejection of a counteroffer is the action of the applicant, not the lender. Overall, he finds that, particularly using corrected data in which counteroffers are dropped, the effects of race are not statistically significant.

The Boston FED study used a single minority dummy variable to test for the effect of race on rejection in equation (1). Some critics have argued that underwriting is much more complex. Horne (1997) reports that the FDIC examiners found evidence in loan files that underwriting decisions are very complex. For example, in the presence of a cosigner, many characteristics of the applicant may be far less relevant. For those with poor credit history, LTV may be the dominant underwriting variable regardless of income and payment to income ratios. Bostic (1996) tested a specification in which various underwriting variables were interacted with a race dummy to see if it appeared that lenders reacted differently across racial groups to the same underwriting variable. He found that the sign of the racial interaction term varied with evidence that, for some variables, being minority was an advantage while for others it was a disadvantage.

Regardless of the concerns of critics, the conventional approach to examination of lenders, by bank regulators, internal bank fair lending divisions and plaintiffs in fair lending cases, inevitably includes estimates of equation (1) using the loan-level information in the institution’s internal databases. Automated underwriting has improved the quality of this data and increased reliance on it in the underwriting process. Just as the banks voluntarily participated in the collection of the data which resulted in the Boston FED study, since then the evolution of bank lending procedures has made it easier to conduct fair lending examinations using estimates of rejection equations as the initial step of the process. Of course, one might imagine that use of automated underwriting would make it unlikely that race could have a statistically significant relation to rejection. However, this has not proved to be the case for reasons that will be apparent later in this report.

III.A.2 Testing for Discrimination Based on Mortgage Pricing Equations

Testing for discrimination in pricing of various forms of consumer credit is not new. The rise of risk-based pricing of mortgage products has increased interest in using statistical models of the price of mortgage credit to make inferences about possible discrimination in lending. More recently amendments to the HMDA reporting requirements required lenders to report the difference between the APR and the comparable maturity Treasury security for all “higher priced loans” exceeding a yield-spread premium.⁸ All this, along with the increasing availability of low cost data from automated underwriting systems has made the estimation of mortgage pricing equations cheap and popular for those wishing to implement a simple test for discrimination.

In an article in the Summer 2005 Federal Reserve Bulletin discussing the expanded HMDA data, Avery and Canner (2005) include a text box that reviews the factors that influence loan pricing. Written as a loan pricing equation, the discussion implies that APR is determined by:

$$\text{APR} = F(i, R, E, S, D, N, C) \quad (2)$$

where: i is the cost of funds which varies with the expected duration of the loan; R reflects the two principal elements of risk in mortgage lending, credit risk and prepayment risk; E is overhead expenses of preparing and processing the loan documents; S is servicing cost which vary with loan type and amount; D reflects discretionary pricing by loan officers; N measures the effects of negotiation by applicants and C is the delivery channel through which the mortgage is acquired. Based on this discussion, it is clear that most of the factors that determine the pricing of individual loans are not available in the expanded HMDA data.⁹

Equation (2) also indicates that, even if expanded data from the loan files of the mortgage originator is available, measuring all the important variables that relate to the loan pricing decision is difficult indeed. While the cost of funds can presumably be observed from market interest rates at the time of endorsement, for loans that float rather than lock in immediately, the cost of funds is ambiguous. Presumably the most important variables that predict credit and prepayment risk are collected by the lender. However, the way they are measured and their effect certainly vary across lenders and even across loan products for given lenders. Furthermore, in what will be the central point made in this essay, the loan terms included in R are not only causes of APR, they are caused by APR. For example, while the likelihood of prepayment may cause higher APR, it is equally true that higher APR makes prepayment more likely. Clearly APR is not only caused by prepayment risk, it is equally true that prepayment risk causes APR.

The discussion of overhead expenses in the text box deserves particular attention. Overhead expenses vary with the characteristics of the applicant. Individuals with poorly documented income, wealth, employment and credit history place additional costs on the loan officer and underwriter. Of course, in the competed loan application, the difficulty of preparation and verification of the information is often not reflected. Servicing costs per dollar of loan amount tend to vary inversely with the duration of the loan and are very high for loans where the probability of delinquency, default and foreclosure are elevated. As the loan amount increases, overhead expenses and servicing costs per dollar of loan amount tend to fall, accounting for the finding that the estimated coefficient of loan amount in an APR equation tends to be negative.¹⁰

Historically, many lenders have permitted loan officers to exercise some discretion in loan pricing. Some loan officers use the low service, high volume, low price approach. They tend to be differentially attractive to applicants who can fill out their own loan forms and document all information required by the underwriter. Other loan officers use the high service, low volume, high price approach and

serve applicants who have difficulty with loan forms and/or documenting their income, employment, wealth or credit history. Avery and Canner (2005) identify two other sources of variation in APR that are related to discretion. One is negotiation, in which applicants use the discretion available to the loan officer to bargain the APR down under threat of dealing with another lender. To the extent that the applicant appears qualified for mortgage credit, such tactics may lead loan officers to use their discretion to cut APR. The other is the effect of delivery channel. Lenders routinely report, for HMDA purposes, loans taken in through very diverse processes. In some cases, the loan officer is not an employee of the lender and the underwriting may even be done by individuals not in the employ of the lender. Even when lenders are dealing with their own employees, the cost structure of serving customers through a loan officer sitting in a bank branch and providing all manner of consumer services and one operating from a remote location where the only activity is taking mortgage applications electronically are very different.

One final cause of variation in APR not noted in Avery and Canner is the effect of applicants failing to lock their rate, i.e., “float.” The APR may be determined at or near the date of application, the date of loan approval, or up to (usually) three days before closing of the loan. This depends, largely on the preferences of the borrower. From the point of view of the loan officer, applicants requesting a quotation at application, before the loan officer has done any processing, are very different than those whose paperwork has been processed and approved. For those borrowers who float until closing, the lending process generally provides a formula for determining the interest rate based on lending conditions about three days prior to closing. This formula is not determined by the loan officer and is generally different than that for borrowers who do not float. Thus floating introduces a problem because there is usually no way of knowing what time path of possible interest rates were quoted to the borrower between application and closing.

In spite of the cautions noted when the expanded HMDA was released by the Board of Governors, some studies have used this data alone and estimated single-equation models like (2) with demographic variables added to measure the effects of borrower characteristics on mortgage pricing. Not surprisingly, with so many missing variables, positive and significant effects of borrower ethnicity and even gender have been reported.¹¹ Estimates of equation (2) have also been used in litigation and in testing for fair lending problems by examiners and lenders themselves.

As noted above, there are multiple reasons for thinking that the structure of mortgage pricing varies by loan channel, particularly by the difference in prime and nonprime channels. Put another way, this means that, even if all the variables in the $F(i, R, E, S, D, N, C)$ function were available for statistical analysis, the functional form would vary by channel, i.e. there would be a $F_{\text{Prime}}(\cdot)$ for prime mortgages and a $F_{\text{Subprime}}(\cdot)$ for subprime mortgages. Courchane (2007) reported estimates of such a model which allows this variation and has done this with a dataset that includes information from loan files that goes well beyond HMDA. Her model includes a selection equation in which applicants are more or less likely to apply for subprime mortgages, and then estimates the $F_{\text{Prime}}(\cdot)$ and $F_{\text{Subprime}}(\cdot)$ functions

with observations weighted inversely by the probability that the mortgage type was the one chosen. The results provide three very important insights beyond the application of the sample selection technique to choice of mortgage channel. First, prime and subprime APR equations are different. Second, the unadjusted difference in subprime use between white non-Hispanic and African-American borrowers is 28 percentage points but falls to only 0.7 percentage points once omitted variables are added to the model. Third, differences in prime and subprime APRs between white non-Hispanic and African-American borrowers fall from 65 and 59 basis points in subprime and prime loans using unadjusted data to 10 and 8 basis points in the full model estimates. Similar findings apply to APR gaps with Hispanic borrowers. Overall, these results suggest that the combination of omitted variable bias and failure to account for differences in lending channel tend to produce very false impressions of the APR differentials from estimates of an APR price model like equation (2). This is not surprising once the large number of variables that should enter estimates of equation (2) is compared to what is available from HMDA or even from moderate enhancements to HMDA data. This statistical exercise makes two points clear. First, current HMDA data are completely inadequate to characterize the loan pricing decision. Second, substantial additions to HMDA data would be necessary to include the major factors determining pricing. However, later in this paper, I will demonstrate why any estimates of APR equations such as (2), even with data enhancements and an estimator allowing for selection by channel, will produce biased and inconsistent results of the relation between borrower demographic characteristics and the lender's pricing policies.

III.A.3 Testing for Discrimination Based on Mortgage Default Equations

The theory of the economics of discrimination states that differential treatment discrimination should be reflected in performance equations. Based on this argument, tests have been designed to test for discrimination in a number of areas outside credit markets. For example, tests for discrimination in selection for professional sports teams have used the relation between performance characteristics of the weakest African-American player selected for the team and the weakest white player. If the weakest African-American player performs better than the weakest white player, this indicates discrimination.¹² Recently Mixon and Trevino (2004) have implemented a test for discrimination in firing head coaches in college football. Their statistical model considers the performance of the team since the coach was hired and models the time until the coach is either dismissed or leaves due to a voluntary separation. The negative and significant estimated coefficient for race was used to argue that, conditional on hiring a minority coach, colleges were more reluctant to fire minority coaches. Madden (2004) tested for discrimination in hiring National Football League coaches by estimating a model of team winning percentage including player and payroll information and adding a variable indicating the race of the coach. A positive and significant partial effect of African-American head coaches on winning percentage was found and taken as an indication of discrimination in hiring of African-American head coaches.

Transferring this testing technique to mortgage lending requires that some measure of loan performance be related to both loan characteristics and demographic characteristics of the

borrower. Profitability of the loan would be ideal but the literature has used default or default loss because profitability is difficult to measure. Equal treatment requires that applicants who are equally creditworthy be treated equally by lenders. If African-American applicants are more likely to be rejected than equally credit worthy whites then default rates should be higher among whites, holding all criteria of the underwriting process constant. Equation (3) has the general form of such an ex-ante default equation:

$$P(D) = f(F, R, T, C) \quad (3)$$

where $P(D)$ is the probability of default, F is a vector of variables reflecting financial capacity of the applicant, R is variables measuring risks of default in the local housing market, T is a vector indicating the loan terms and C is indicators of personal characteristics of the borrower, including race. Equation (3) is called an ex-ante default equation because the independent variables, represented by F , R , T and C , measure characteristics of the loan, borrower, and collateral that are observed at application. Later in this essay there is a discussion of ex-post default in which the condition of the loan over time since endorsement is considered. Similarities between equations (3) and (1) are not coincidental. A primary reason for loan denial is the expectation of costs imposed by delinquency, default and /or foreclosure. Indeed, the requirement that a business reason be provided for non-discriminatory rejection, means that the denial equation should be quite close to the default equation. As with equation (1), demographic characteristics are added to equation (3) in order to implement the test for discrimination. Evidence of discrimination is provided by a negative and significant estimated coefficient for a prohibited variable in the default equation.

Estimates of equation (3), performed using data on FHA-insured mortgages by Berkovec, Canner, Gabriel and Hannan (1994), found that the estimated coefficient of the variable indicating the borrower was African-American was positive and significant. Rather than indicating discrimination against African-Americans, this result indicates relatively favorable treatment in the approval process. It is important to note that these tests were performed on FHA-insured mortgages. While these mortgages were directly endorsed by private lenders, the lending is subject to more government oversight than conventional mortgage lending and hence critics of the industry argue that the results do not reflect the industry as a whole.

Just as the Boston FED rejection equations with their positive African-American effect had aroused considerable comment, the estimates of default and default loss equations with their positive African-American effect prompted considerable reaction from the research community. Of course, the positive coefficient in the single equation rejection model was taken as evidence of discrimination while the positive coefficient in the default equation indicated non-discrimination or, if anything, discrimination in favor of African-American borrowers. Indeed, most of volume 2, number 1 (1966) *Cityscape* journal published by HUD was devoted to a discussion of the merits and flaws of single equation default models

as indicators of discrimination. Some of the criticisms are standard, such as the problem of omitted variables bias which works just the opposite in rejection and default equations. Others concern the possibility that there is discrimination in the FHA foreclosure or mortgage servicing process.

Subsequently, Berkovec (1998) developed a more subtle version of the default equation test in which he argued that lenders in more concentrated markets have greater ability to act in a discriminatory fashion. His test for discrimination then involved a measure of lender concentration interacted with the racial type of the borrower in a default equation. The discrimination test then became whether minorities in more concentrated lending markets had lower default rates than those in more competitive mortgage markets where discrimination would be more difficult. The findings showed no significant effect of market concentration. This test answered much of the criticism directed at the initial findings but, as will be clear later in this essay, it still suffers from the same biases that affect all single-equation tests for discrimination in mortgage lending.

III.B Conventional Approaches to Measurement of Credit Risk

The recent financial crisis revealed many shortcomings in the mortgage market. One was that default and default loss models woefully underestimated credit losses. This section will demonstrate that the poor performance of statistical models of mortgage default was the natural product of the fact that conventional statistical models ignored economic theory and assumed that loan terms were chosen without regard to the probability of future default. Paradoxically, the same failure to apply economic theory and model loan terms as endogenous to mortgage rejection and pricing also produced a failure in mortgage default models that became most acute as expected house price volatility and mortgage default rates both rose.

A number of innovations of the last 30 years have led to dramatic changes in mortgage underwriting and pricing. Standardized loan applications promoted by the need to securitize mortgage credit have made high quality loan-level data from application through final termination available to researchers, lenders and investors in mortgages.¹³ Information on the performance of these loans was then collected and used in mortgage loan intelligence models. Data on the performance of individual loans over time allowed analysts and investors to follow subprime mortgage pools on a monthly basis. The market for derivative securities based on these mortgages raised demands for conditional forecasts of cash flows due to mortgage terminations and delinquencies in order to price the securities.

Automated underwriting based on default loss modeling allowed lenders to economize on underwriting costs and made growth of the subprime mortgage lending industry possible. The business model for subprime lending involved underwriting large numbers of problematic applicants, high rejection rates and risk-based pricing. Given the high probability of rejection, applicants were understandably unwilling to pay application fees that would cover even a small fraction of the normal underwriting cost on a prime mortgage. Accordingly, underwriting costs had to be kept low for this type of lending

to grow and this was possible as underwriting became increasingly automated.¹⁴ In order for this underwriting to be automated, statistical models of the relation between applicant characteristics, loan terms, market conditions and the probabilities of default such as that given by equation (3) were necessary.

Two types of statistical models have been used to measure credit risk. The first is an ex-ante default or default loss model having the form of equation (3). These models are based on “seasoned” mortgages relying on von Furstenberg’s (1969) classic result that the probability of default loss falls drastically with time since endorsement. Ex-ante default models used to support mortgage lending do not include demographic characteristics of the borrower and have been the object of some scrutiny in order to avoid adverse impact discrimination. Indeed, the increasing use of these models has allowed discussion of adverse impact discrimination to gain a solid statistical basis as the business necessity of using particular variables can be weighed against the impact on protected groups. Furthermore, use of these standard models eliminates the chances for statistical discrimination as lending decisions are made on objective criteria and the necessity for direct contact between the applicant and loan officer is eliminated.¹⁵

The second type of model is designed to estimate the cash flow from mortgages and may be used either before or after endorsement. It is an ex-post model in that it includes both variables reflecting conditions at application and those reflecting the evolving conditions of the mortgage and housing market. Measurement of cash flow requires modeling both expected default and prepayment terminations over the remaining term of the mortgage. The very high early prepayment rates for subprime mortgages made this type of modeling essential if mortgages remaining in a seasoned pool were to be priced.¹⁶ The models of credit and prepayment risk that can be used for mortgage pricing have the general form:

$$\begin{aligned} \Pr(D) &= f(B_{Dt}, F, E_t, T, C) & \mathbf{(4)} \\ \Pr(P) &= g(B_{Pt}, F, E_r, T, C) \end{aligned}$$

where P(D) is the probability of default, P(P) is the probability of prepayment, B_{Dt} and B_{Pt} (for Black-Sholes option model) are vectors of variables measuring the value of the option to default and prepay respectively, F is a vector of variables reflecting the financial capacity of the applicant at endorsement, E_t is variables measuring economic conditions which vary over time, T is a vector indicating the loan characteristics and C is indicators of personal characteristics of the applicant, including race. Note that some of these variables are constant characteristics of the transaction or applicant and others are time varying.

Finding an appropriate statistical estimator for the equation system in (4) is complicated because prepayment and default are the result of failure processes in which mortgages either survive another time period or fail for one of two reasons. Thus the mortgages surviving into the second year are

fundamentally different than the population of initially endorsed mortgages because they failed to either prepay or default in the first year. In statistical jargon, surviving mortgages have been selected to survive and hence are systematically different than the initial population. Estimating (4) in a fashion that allows for this selection process and also measures the economically correct value of the options to default and prepay have provided a major research challenge for many years.

A series of papers by Kau, Keenan, Muller and Epperson¹⁷ provided theoretical models that established the interdependence of prepayment and default options and hence the desirability of estimating the system of two equations in (4) jointly. Ambrose, Buttimer and Capone (1997) made further improvement in modeling prepayment and default options, by explicitly introducing into the option-pricing framework the delay of foreclosure and the concept that the decision to stop making payments is determined by expected future values of the property. Deng, Quigley and Van Order (2000) applied the Cox proportional hazards model with group duration data to analyze empirically residential mortgage prepayment and default behavior using micro data on the joint choices of individuals. Their econometric model of the competing risks of mortgage termination by prepayment and default accounting for borrowers' heterogeneity has become a workhorse model for empirical analysis of mortgage terminations. Another alternative to duration models used in empirical literature on mortgage terminations is multinomial logit models with restructured event history.¹⁸

Estimates of the model in (4) have yielded a number of interesting results. Capozza and Thomson (2005) report that because they have a longer period of delinquency, subprime loans tend to inflict larger losses than prime loans. Danis and Pennington-Cross (2005) found that delinquency and default vary with changing local economic conditions, housing market conditions, credit scores and loan characteristics. Ho and Pennington-Cross (2006) report evidence that subprime loans terminate faster than loans in the prime market, and the hybrid loans terminate at higher rates than fixed loans. Pennington-Cross (2006) argued that foreclosures on subprime mortgages are affected by many factors including contemporaneous housing market conditions, the prior performance of the loan (prior delinquency) and the state-level legal environment. Quercia, Stegman and Davis (2005) found that loans with prepayment penalties and balloon payment requirements have a significantly higher mortgage foreclosure risks, controlling for other risk factors, such as borrowers' credit history, loans' characteristics and purpose, housing type and state-level macroeconomic fundamentals. In subsequent sections of this paper, the reasons for these results will be apparent.

Rose (2008) found that the relation between loan terms and the probability of foreclosure varies significantly for subprime refinances and home purchase mortgages, and that within these categories there are further differences for fixed and adjustable-rate mortgages. This last finding is particularly consequential because it suggests that the underwriting and pricing of prime and subprime mortgages in general and even by detailed type of mortgage, should be based on different models. Put another way, the automated underwriting scheme used to deny and price a mortgage application should differ for prime and subprime mortgages, and even by type of prime and subprime mortgage.

This brief review of the academic literature modeling credit and prepayment risk on mortgages is presented to illustrate the type of work and results that have been obtained. Even more consequential were the models estimated privately and used to price mortgages and subsequently pools of mortgage backed securities. Investors placed considerable confidence in the predictions produced by such models in spite of the statistical problems that are discussed in the next section of this report.

IV. Problems with Conventional Approaches to Measurement of Mortgage Market Discrimination and Credit Risk

To be clear, the statistical problems arising from a neglect of economic theory that have been discussed above not only call into question findings of lending discrimination, they also are a warning regarding the potential robustness of any single-equation mortgage credit model. Having reviewed conventional methods used to evaluate mortgage market discrimination and credit risk, this report next considers important statistical problems that question the credibility of the statistical estimates resulting from these models. The arguments made here are not new. Indeed, they predate virtually all of the literature referenced thus far. Furthermore, the arguments imply that the statistical estimates of the parameters of equations (1), (2), (3) and (4) are biased. The problems do not disappear with larger or more comprehensive data sets because they arise from a faulty model and statistical method. In some cases, it is relatively easy to show the nature and direction of these biases in estimates. For example, the biases in estimates of rejection equations tend to produce false positive indications of default and the biases in default equations tend to underestimate default losses. Some of the recent problems in performance of mortgage markets may be related to these biased statistical estimates of discrimination and credit risk.

How did these empirical techniques for measuring mortgage market discrimination and credit risk that produce biased and inconsistent results originate? Why do they persist? What is the nature of the problems with specific statistical approaches? The next subsection will discuss some standard statistical problems that arise when trying to estimate conventional models of discrimination and credit risk in mortgage markets. This is followed by a two sections that first establish the lack of theory necessary to support these conventional approaches and the minimal level of theory needed to support valid statistical estimates. The last two subsections provide a specific criticism of the three conventional tests for discrimination and for previous approaches to modeling credit and prepayment risk based on their lack of theoretical and statistical support.

IV.A. Sources of Bias in Conventional Statistical Approaches

All of the models discussed above involve statistical tests that relate some random “dependent” variable, rejection, pricing, default or prepayment, which is the outcome of the mortgage transaction process to other “independent” variables that characterize the applicant, property, mortgage terms and economic environment in which the transaction occurs. In applying statistical methods to data generated by controlled experiments, it is possible to change one independent variable at a time, hold other independent variables constant and observe the change in the dependent variable. The causal relation between the independent variable and dependent variable is insured by the experimental design and other variables are held constant by that same design.

Unfortunately, mortgage data is not experimental.¹⁹ Applicants choose where they will apply and the type of mortgage terms that they will accept strategically. Obviously, lenders behave strategically also. Furthermore, some characteristics of the transaction are unobserved in the data while others are measured with error. Consider a variable like the down payment. Is this determined by the borrower based on wealth and portfolio considerations? Is it determined by the loan officer in response to underwriting criteria that limit LTV or make rejection likely? Perhaps, it is determined by the altruism of family members? Whatever the case, the important statistical point is that down payment and other aspects of the mortgage transaction cannot be controlled and hence must be modeled and well understood before doing statistical analysis.

The nature of the statistical problems encountered when empirical work goes forward in the absence of theory is easily demonstrated. Indeed, the points made here are not new. They date at least to Barth, Cordes and Yezer (1981) and Maddala and Trost (1982) and have been reaffirmed in a host of subsequent studies. Single equation models of mortgage rejection, pricing, default and /or default loss all relate mortgage outcomes to a variety of “causal” variables. The dominant “causal” factors are usually loan terms: amount financed, LTV, monthly payment, payment-to-income ratios (front and back end), cosigner, etc. In a single-equation model, the assumption being made in the statistical analysis, often made implicitly without discussion, is that the mortgage outcome variables have no role in causing the loan terms. For example, single-equation models must assume that applicants have no knowledge of the relation between the loan terms that they request and the probability of rejection. If the probability of rejection is causing applicants to increase their down payment or otherwise modify loan terms, then the causal assumption justifying single-equation models is violated and they will produce biased and inconsistent estimates of the true rejection equation of the lender. In the technical discussion below, it can be demonstrated that rejection equations tend to produce false positive indications of discrimination against disadvantaged minorities while single-equation default models tend to produce false negative indications. Thus the contradictory results actually reported in the single-equation literature are explained by the bias in single-equation models.

It is possible to formalize the arguments just made regarding the relation between theory and much of the current empirical work on mortgage markets. Consider a simple relation between one random variable, R, and three other random variables L, Y and Z. In this case, R can be rejection of a mortgage application, L the LTV and Y and Z the variables that indicate the creditworthiness of the applicant. Following equation (1) above, we might write the relation among R, L, Y and Z as:

$$R = a + bL + cY + dZ + u \quad (5)$$

Here a, b, c and d are parameters of the rejection relation that are unknown and need to be estimated statistically and u is a random variable or error term reflecting the parts of the rejection decision not captured by the three included variables or perhaps measurement error.²⁰

Statistical estimation of equation (5) using ordinary least squares requires a number of assumptions but the one that is of most concern here is that the expected value of the error term u cannot be correlated with L, Y or Z. This may be stated as $E(u|L) = 0$, the expected value of error term u given L equals 0, or $r_{uL} = 0$, the correlation between u and L equals 0. Similarly unbiased estimates of c and d assume that $E(u|Y) = E(u|Z) = 0$ or $r_{uY} = r_{uZ} = 0$. To see why these properties are important, recognize that equation 5 states that R can be predicted by L, Y and Z, so that the expected value of R, given L, Y and Z is: $E(R|L,Y,Z) = a + bL + cY + dZ$. Clearly this can only be true if $E(u) = 0$, which requires that $E(u|L) = E(u|Y) = E(u|Z) = 0$.

IV.A.1 Omitted Variable Bias

Now consider what happens to statistical estimates of (5) when data on Z are not available. This can give rise to omitted variable bias through the following mechanism: dZ will now be part of the error term of the estimates, i.e. the regression error will be dZ + u. Now statistical estimates of (5) will include only the two observable variables, L and Y and the estimator assumes that $E(dZ + u|L) = E(dZ + u|Y) = 0$ or that $r_{L(dZ+u)} = r_{Y(dZ+u)} = 0$. If this is true then the expected value of R, given L and Y will be $E(R|L, Y) = a + bL + cY$. However, it is unlikely that L, Y, and Z are uncorrelated because loan terms and applicant characteristics tend to be related. Let's say that Y and Z are correlated and that, the estimated coefficient of an ordinary least squares regression of Y on Z would be α . Now the expectation of R given L, Y is $E(R|L, Y) = a + bL + cY + d\alpha Y = a + bL + (c + d\alpha)Y$ and the ordinary least squares regression estimator will produce an estimate of c equal to $(c + d\alpha)$. Clearly this estimate of c is biased by the term $d\alpha$ and thus is the classic case of omitted variable bias. In tests for discrimination, imagine that Y is a variable indicating minority status of the applicant and Z is an indicator of creditworthiness. It follows that $d < 0$ because increased creditworthiness lowers the probability of rejection. If minorities are generally less creditworthy, then $\alpha < 0$. It follows that the product $d\alpha > 0$ and the omitted variable bias is positive so $c + d\alpha > c$ and statistical estimates of (5) will tend to produce positive indications of discrimination even if $c = 0$.

Omitted variable bias produces similar results in estimates of APR equations and default equations. In the case of APR equations, the bias produces false indications that minority borrowers pay more for credit which is an indication of differential treatment discrimination. However, in the case of default equations, the omitted variable bias also raises the estimated coefficient of the minority status variable making it appear that minority borrowers are more likely to default. Paradoxically, omitted variable bias in the rejection and APR relations tends to produce false positive indications of discrimination and the same bias tends to produce a false negative indication of non-discrimination in a default equation.

It is worth noting that conventional estimates of denial and APR equations reviewed above often have positive estimated minority coefficients, taken to indicate discrimination, while default equations have positive estimated minority coefficients, interpreted as indicating non-discrimination. Thus, if there are problems of omitted variable bias in conventional statistical estimates of rejection (equation (1)), APR (equation (2)) and default (equation (3)), this could explain the paradoxical tendency to find positive indications of differential treatment discrimination in the first two cases and evidence of non-discrimination in default equation estimates. This also explains why the estimated coefficient of minority status in rejection and APR equations is large when only HMDA data are used and falls as additional variables are added to the regression estimates. Such results are a classic indication of omitted variable bias. While there is an indication of omitted variable bias problems when addition of more variables changes the estimated coefficients of the other independent variables, there is no way to eliminate the bias without further efforts at data collection.

Does this mean that more data on additional variables related to the loan transaction will produce unbiased estimates? Unfortunately, omitted variable bias is only one of the sources of bias in the estimated coefficients of rejection, APR and default equations and these other problems will not be eliminated by additional data collection.

IV.A.2 Other Problems

In addition to omitted variable bias, there are other statistical problems in conventional attempts to estimate equation (5). Note that the equation assumes that the relation between rejection and L, Y and Z is linear or that the effect on the expectation of R of unit change in L, Y or Z is exactly b, c or d. It may be that the effect of Z on R is non-linear, so that the true model is $R = a + bL + cY + dZ + \partial Z^2 + u$. Now the error term of the regression in equation (5) is $\partial Z^2 + u$ and clearly this error term is positively correlated with Z so that $E(\partial Z^2 + u|Z) > 0$. Once again, the ordinary least squares estimates of d will be biased upward. The problem with this type of specification error is that there is very little to guide researchers in determining the functional form of the rejection, APR or default equations. Ordinarily researchers test many alternative specifications in order to guard against the possibility of specification error but this is difficult when theory provides so few restrictions on functional form.

We now have the necessary intellectual ingredients to consider another major problem in conventional statistical tests for discrimination and credit risk in mortgage lending that arises from the failure to use economic theory in modeling the mortgage application and approval process. That problem is simultaneous equations bias which arises because economic theory implies that some of the “independent” variables are not actually independent. Considering the rejection equation (5) above, it is logical to write this as part of a two equation system where:

$$R = a + bL + cY + dZ + u_R \quad (5)$$

$$L = e + fR + gY + hZ + u_L \quad (6)$$

In equation (6), the LTV, L , is written as a function of the probability of rejection, R , and the independent variables Y and Z . Note that subscripts attach the error terms to the R and L equations. Why does the probability of rejection enter the LTV equation (6)? The theory behind this will be discussed more formally later in this section, but the basic economic argument is that applicants increase their down payment, i.e. lower the LTV, as the probability of rejection rises, i.e. $f < 0$. Applicants are motivated by the desire to avoid rejection and one of the primary choices that they can make to insure approval is to lower the LTV sufficiently. If L is determined by (6), then it is clear that $E(u_R|L) \neq 0$ because (6) tells us that higher R is associated with lower L , and hence $r_{LuR} < 0$ or $E(u_R|L) < 0$. This means that ordinary least squares or other single-equation model estimates of the parameters, a , b , c and d , in equation (5) will be biased and inconsistent.

The likely nature of the bias and its implications for statistical models of discrimination and credit risk is more complex than the case of omitted variables bias and will be considered carefully in a subsequent subsection. However, the general finding in the case of rejection equations is that conventional single-equation models tend to produce false positive indications of discrimination. The reason for this is that households with more resources are better able to avoid rejection at the margin by supplying additional down payment, getting cosigners, etc. and even having their applications reconsidered in light of these credit enhancements. To the extent that minority applicants have less access to such additional resources, they are less able to avoid rejection and /or high APR. In a sense, this result is similar to any other market result. Individuals with greater personal resources are generally better able to avoid adverse outcomes.

There is a standard remedy to the problem of identifying the parameters of equation (5) and that is to find identifying information in the form of variables that belong in equation (6) but can be excluded from equation (5). These variables can be used to identify movements in L that are independent of u_R . Unfortunately, such variables are difficult to find in research on mortgage credit because the loan-level data collected by lenders and used in the estimation is specifically designed to support the underwriting process. That is, lenders are motivated to collect information on borrower creditworthiness

and collateral value only insofar as these help to evaluate the risk of lending. Indeed, any lender who collected information from applicants that was not designed to aid the underwriting process might well be suspected of using that information for some discriminatory purpose. Thus the problem of simultaneous equations bias persists precisely because the basis for modeling lending behavior is examination of the underwriting process.

The example of simultaneous equations bias given here was the LTV. However, any variable over which the applicant has some control and which can be used to influence the underwriting decision has the same characteristics. As noted many years ago by Barth, Cordes and Yezer (1980), probability of rejection, probability of future default and prepayment, loan amount, value of collateral, monthly payment, points, interest rate and prepayment penalties are surely jointly determined endogenous variables. This means that, rather than a one- or two-equation model of mortgage lending, a model with many equations must be specified and estimated if unbiased estimates of the parameters are to be obtained. Compared to this standard, conventional efforts at estimating one- or two-equation models are completely inadequate and are likely to generate false positive indications of discrimination in rejection or APR equations.

IV.B Lack of Theoretical Support for Conventional Statistical Approaches

The previous section demonstrated that, when estimating models using non-experimental data, strong conditions on the relation between the error term and the “independent” variables must hold. In the case of mortgage market transactions, where the independent variables that determine creditworthiness are many and may lack standard measures and loan terms are selected by the applicant, some care must be taken to avoid problems of omitted variable bias and simultaneity. These statistical problems are not uncommon in economic models. They usually prompt careful theoretical modeling of the processes that create the observed transactions followed by a section on stochastic specification that anticipates problems like omitted variable bias and jointly determined variables. Detailed arguments are then made to justify the statistical approach used to avoid these problems.

Even a cursory reading of the literature reviewed in section III on “conventional” approaches to measuring discrimination and credit risk in mortgage lending reveals that there is no formal economic theory justifying the equations being estimated and there is no stochastic specification that demonstrates that the estimates are unbiased. The lack of attention to theory may partially be blamed on the extreme complexity of the mortgage transaction itself. It is also very convenient because, even a cursory examination of available models of the mortgage application process demonstrates that the assumptions necessary to support the use of conventional statistical techniques cannot be justified.

Consider, for example, what passes for a theory of the mortgage lending process in the Boston FED study. There is a section that says it is a model of mortgage lending. This section says that lenders

maximize expected profit from the loan and that the “primary task facing the lender is avoiding default and any associated losses.” There is no discussion of the motivation of the applicant and the rejection equation (1) is described in terms of the lender’s decision alone.²¹ In reviewing the literature on discrimination some years later, Ladd (1998) not only fails to present any theory of the mortgage lending process, she fails to note that one might be needed in view of the literature on simultaneous equations bias in single-equation mortgage rejection models. Overall the standard practice in studies using conventional statistical models of mortgage discrimination, pricing and even default is to present no theoretical model of borrower-lender interaction and to pretend that the transaction is based on decisions of either the borrower or the lender acting in isolation.

IV.C Theoretical Support for Valid Statistical Models of Discrimination and Credit Risk in Mortgage Lending

Students of economics are taught, usually early in their undergraduate training, that empirical testing of economic models should only proceed after careful development of a theoretical model relating the variables being studied. The reason for this training is that empirical testing in economics is not based on experiments in which a single experimental variable is observed while all other variables are controlled. The usual model of experimental science has a single random variable whose outcome is the unknown consequence of variation in other control variables.

Non-experimental testing, particularly using data from mortgage markets, has no control variables.²² Undergraduate students are warned that, in the absence of careful theoretical models that place strong restrictions on the relations among economic variables, no strong statements about cause and effect can be made. Perhaps the most famous example of an empirical relation that was developed without theory and used for policy purposes in economics is the Phillips Curve. In 1958, William Phillips published a paper entitled “The Relationship between Unemployment and the Rate of Change of Money Wages in the United Kingdom 1861–1957.” This was an empirical paper, not based on economic theory, which said that the two random variables, unemployment and the rate of change in nominal wages, were inversely related. In 1960, future Nobel laureates Robert Solow and Paul Samuelson produced a similar empirical paper for the United States, again without theoretical support for the relation between unemployment and wage change. These papers, along with other similar work, were highly influential in guiding public policy and creating expectations that government policy could eliminate business cycles.

Of course, subsequent development of theoretical models of worker behavior based on rational expectations has shown that there is a natural rate of unemployment and that the long run relation between the unemployment rate and either inflation or wage change is vertical. For an extended period of time even after publication of the rational expectations critique of the Phillips Curve, however, very unfortunate economic policy decisions were made based on the notion that the government could control the long run unemployment rate by raising inflation. This is a classic example of the

problems that are encountered when statistical relations among economic variables are taken seriously without a theoretical model of the relation among the variables to support application of the empirical technique. The major contention of this essay, is that, like the empirical Phillips curve, current statistical techniques for measuring discrimination and credit risk in mortgage markets lack a theoretical basis. Current statistical tests for discrimination and credit risk lack theoretical support and, like the empirical Phillips Curve, have produced false and misleading results that have been incorporated in public policy and private practice.

Unlike the relation between unemployment and inflation or wage change for which satisfactory theoretical models have been produced and tested, there is no fully developed theoretical model of the mortgage lending process. However, even the partial theoretical models that are in the literature are sufficient to demonstrate that current statistical approaches, particularly the single-equation models, are invalid.

The lack of a comprehensive theoretical model is due to the complexity of the lending process and the mortgage product. In this subsection, the issue of prepayment risk and the role of points and penalties in controlling this risk are ignored and the focus is entirely on credit risk.²³ Even with this simplification, no comprehensive model of the mortgage lending process has been developed.

The mortgage market is an example of a transaction in which one agent, the applicant, has superior knowledge regarding a contract that imposes a contingent liability on the other agent, the lender. Characteristics of transactions in this type of market were first considered for the specific case of insurance by Rothschild and Stiglitz (1976). Their important finding is that equilibrium in such markets may not exist and, if it does exist, it will involve suppliers offering a schedule of prices and limits on quantity to applicants. The mechanism for determining equilibrium follows fairly simple logic. Applicants differ in likelihood of making an insurance claim. A pooling equilibrium in which a single price for insurance clears the market at zero expected profit is not an equilibrium because a clever firm can find an insurance contract that contains a price and level of insurance that is differentially attractive to the lower risks. When these low risks are drawn from the insurance pool, profits of firms are negative and they are led to offer specialized contracts that create a separating equilibrium between the high and low risk borrowers.

Brueckner (1994a) applied this same reasoning to a mortgage market where applicants differ in cost of default and lenders find a separating equilibrium in which borrowers who can default at low cost choose higher LTV loans at higher APRs and are much more likely to default than those with high default costs who are offered contracts with low LTV and APR. Thus whatever initial differences in ultimate default would be expected, based on initial differences in default costs, which are magnified by the separating equilibrium. The important point is that the relation between default and LTV and /or APR that might be expected if all borrowers were served in a pooling equilibrium is not observed in non-experimental data on actual loan transactions. Instead, the separating equilibrium introduces

an additional association between LTV and APR in that borrowers with low default cost will tend to choose high LTV and APR. Furthermore, the borrower characteristics that are associated with low default cost are not observed by the lender and hence will be omitted variables in any empirical model of default, particularly those characterized by the model in equation (3).

More recently, Nichols, Pennington-Cross and Yezer (2004) have shown that for mortgage applicants who differ in creditworthiness in ways that can only be determined by costly underwriting, a pooling equilibrium does not exist in which a single lender serves high and low risk applicants using a single underwriting scheme in zero profit equilibrium. Again the problem is that, in the zero-profit pooling equilibriums, a new entrant can attract the low risk applicants by offering lower APR with higher cost underwriting. The result is a separating equilibrium in which the low risk applicants have low APR, more thorough underwriting and lower rejection rates than the high risk applicants. This implies that higher risk applicants will self select into loan programs with lower application costs, higher APR and higher rejection and default rates. Estimates of rejection, loan pricing or default equations, such as (1 through 4 above) will find that whatever personal characteristics cause applicants to self select into the higher risk category will be positively associated with APR, rejection and default. What is important to note in this theory is that the higher APR, rejection and default experienced by these applicants is due to their self selection into a particular loan program and not due to differential treatment by the lender. Furthermore, the separating equilibrium is forced on lenders by the competitive process, i.e. it is justified by a business purpose.

IV.D Critique of Conventional Models of Mortgage Market Discrimination

Tests for discrimination in mortgage lending, whether based on the sign and significance of a dummy variable indicating minority status in a rejection, APR or default equation have common statistical properties. Based on the arguments made above, none of these single-equation models are adequate to test for discrimination in mortgage lending. Because the Boston FED rejection equation estimates have been given most attention, they will serve as the principal example in this subsection.

The Boston FED model of mortgage rejection follows equation (1) which is repeated here for convenience of exposition:

$$P(D) = f(F, R, L, T, C) \quad (1)$$

where $P(D)$ is the probability of denial, F is a vector of variables reflecting financial capacity of the applicant, R is variables measuring risks of default, L measures of potential loss in default, T is a vector indicating loan terms and C is indicators of personal characteristics of the applicant, including race, that are the object of the test for discrimination.

The discussion of omitted variables makes clear that variables playing a role in indicating creditworthiness which are not included in the regression will lead to false positive indications of discrimination if they are correlated with minority status. For example Zandi (1993) noted that inclusion of a variable reflecting the overall evaluation of the applicant's credit history caused the minority dummy to become non-significant.

Such omitted variable bias is certainly important but it is not the only focus of this critique. The problem with single-equation models of rejection is that many of the "independent" right-hand side variables are jointly determined with the probability of default. Specifically, any elements of the financial capacity of the applicant, risks of default and measures of potential loss, determined by the choices of applicants who wish to avoid rejection, are endogenous variables. This point is well established in the academic literature. As noted in Barth, Cordes and Yezer (1981), Maddala and Trost (1982) and the literature which has applied Rothschild and Stiglitz (1976) to credit markets, the mortgage transaction involves asymmetric information and applicants trade on their informational advantage to make strategic choices. Applicants fully recognize that the probability of rejection is a function of the variables in the loan file and that they may influence these variables to lower rejection probability. The most obvious variables available to the applicant are the loan amount, the value of the property purchased, the monthly payment, term to maturity, points, interest rate, down payment, mortgage insurance and cosigner (if any). These variables are interrelated because the monthly payment is approximately the product of the interest rate and loan amount. Points, down payment, value and loan amount are also related. Thus rejection is a jointly determined variable along with all the other variables chosen by the applicant in order to manipulate the probability of rejection.

Furthermore, the notion of "rejection" by a lender reveals the complex nature of the interactions between applicants and loan officers. It is not uncommon for lenders to decline to accept the initial terms proposed by the applicant but to suggest that other terms such as increased equity or a cosigner, would make the loan acceptable. These terms are then offered to the applicant who may reject them, accept them or propose alternative loan terms. In such cases applicants reject the lender rather than lenders rejecting the applicant. Successive rounds of negotiation may result in ultimate endorsement of the loan or failure to agree. In some cases, loans may be classified as "withdrawn" but generally once there has been initial underwriting done, failure to agree is classified as "rejection" by the lender even if the applicant rejected the last offer made by the loan officer.

The Boston FED study itself recognized that loan terms and ultimate rejection is the result of negotiation between applicants seeking to avoid rejection and loan officers:

"Similarly, if white applicants are more likely than minority applicants to be "coached" when filling out the application, they will have stronger applications than similarly situated minorities. In this case, the ratios and other financial information in the *final* application, which is the focus of this analysis, may find themselves to be the product of differential treatment. This

study does not explore the extent to which coaching occurs...” (Emphasis present in the original, page 43.) Munnell et al (1992):

Hunter and Walker (1996) referred to coaching of non-minority borrowers as the “cultural gap” hypothesis. Horne (1997) documented the fact that lenders make counteroffers to applicants who then must decide whether to accept or reject the terms offered. If the applicant rejects the counteroffer, the application is recorded as a rejection by the lender. Ladd (1998) discusses the negotiation between applicants and lenders in terms of the “thick file” phenomenon:

“The files of white borrowers were likely to end up thicker than those of minority borrowers, and because of that assistance, may have been more likely to be approved.” (pg. 48)

This counteroffer, coaching or thick file behavior, whether related to minority status or not, is direct evidence of joint determination of rejection and the other variables used by lenders in the underwriting process. The simple point is that, if this type of negotiation over loan terms is going on, then what is measured as “rejection” is the outcome of a complex process that should be the object of a formal economic model. Even more important, the causes of rejection cannot be determined by a single-equation econometric model such as that used in the Boston FED report. Put another way, the discussion of “coaching” in the original Boston FED study described in the 1992 working paper is logically inconsistent with the single-equation model that is used to “test” for discrimination in that study.²⁵

Attempts to find variables that are important causes of the endogenous variables controlled by the applicant must discover variables that are not used by lenders in the underwriting process within a dataset collected using information from loan files used in that very process. Consider that loan officers have little incentive to collect detailed information on applicants who are either obviously qualified or unqualified. Horne (1997) documents this in reports from the FDIC examiners who reviewed the loan files used in the Boston FED study. Costly verification is likely only in the case of marginal applications. In cases where underwriters find information in the original application either unverifiable or false, they do not go back and change the application. Accordingly the data recorded from the application for statistical analysis will not reflect the information relied on by the underwriter who uncovered lies in the application.

Also, the amount of information required varies with the nature of the mortgage product. The extreme example is the low or limited documentation loan product in which information on income, assets, employment history, etc. may be recorded in the loan file but do not serve the usual function in credit supply decisions. One purpose of applicant reconsideration is the provision of either information on other sources of income, explanation for previous credit problems or verification of facts in dispute. Thus both the amount and the verification of income, assets, etc. recorded in the loan file are determined, in part, by the probability of rejection and the cost of credit to the applicant. The availability of better loan terms provides applicants with the incentive to provide a more complete or precise accounting for income, assets, liabilities, etc.

Once it is recognized that variables like LTV, monthly payment to income ratio and use of a cosigner belong in the structural rejection equation, then the identification problem created by joint endogeneity of these variables with the rejection rate cannot be solved with the Boston FED data set or with any data set based on loan file information. Thus structural rejection and loan terms equations are not econometrically identified and estimates obtained using single-equation techniques or by attempting to instrument for only one of the jointly endogenous variables are biased and inconsistent. Furthermore, Yezer, Phillips and Trost (1994) have shown that the bias in single-equation rejection estimates can be signed and that the tendency is to produce false positive indications of discrimination. Specifically, they show that, if minorities are financially disadvantaged and tend, other things equal, to make lower down payments, single-equation estimates of the coefficient of a minority dummy variable are biased upward.

Based on the arguments made above, the Boston FED dataset is inadequate to identify a structural rejection or LTV equation. However, it is possible to use reduced-form estimates to either test for bias or to reveal bias in single-equation tests for discrimination in rejection or loan terms. This illustrates the way in which reduced form estimates may be used to test for problems with single-equation models. The specific application of this approach would depend on the nature of the initial single-equation model that was being used to test for discrimination. For expositional purposes, attention is focused on the two most important jointly determined variables, rejection and LTV. The other variables that are usually thought to be jointly endogenous, including all payment-to-debt ratios, term-to-maturity, APR, private insurance and use of a cosigner are ignored. The lender's decision regarding the loan application is given by:

$$R^* = \sum_i \delta_i d_i + \beta_L L + X'\beta + \varepsilon_R \quad (6)$$

where: R^* is unobserved by the econometrician except that, if $R^* > 0$, the application is rejected, $d_i = 1$ if the applicant is a member of the i th demographic group (e.g. married minority, single minority female, etc.), 0 otherwise, L is the LTV, X is an array of exogenous variables (excluding negotiated loan characteristics) and ε_R is an iid error term. As they are coached by or negotiate with the loan officer, applicants observe R^* with error according to $R^{**} = R^* + u$, where u is also an iid error term. We anticipate that $\beta_L > 0$ because higher LTV results in greater chance of rejection.

Borrowers adjust their LTV, L , to achieve a utility-maximizing pair of R^* , L , conditional on their information set R^{**} , according to:

$$L = \sum_i \lambda_i d_i + \alpha_R R^{**} + X'\alpha + \varepsilon_L \quad (7)$$

where it is assumed that $\alpha_R < 0$ as borrowers increase down payments in order to avoid rejection. There are no compelling theoretical arguments providing restrictions on the signs of the estimated coefficients of the demographic dummy variables in equations (6) or (7). The δ_i 's reflect the effects

of differential treatment of applicants based on demographic factors. Indeed, some demographic characteristics, like gender, have generally been dismissed as the basis for lender discrimination while others, like race, have been considered to be significant. Although there is no theoretical reason for demographic variables in the LTV equation, non-zero λ_i 's may arise from differential behavior among demographic groups (for which we lack a persuasive explanation) or from omitted variables correlated with demographic characteristics. Some demographic groups may either differ in their sensitivity to the possibility of rejection or, more importantly, some may have greater resources to bring to the table to avoid rejection.

The equation system (6) and (7) cannot be estimated directly because L and R are jointly determined variables and there is insufficient information to separate them, in econometric terms they are not identified.²⁶ However, structural equations (6) and (7) may be solved algebraically to yield the following reduced-form equations, which can be estimated statistically, for application rejection and LTV respectively:

$$R^* = \sum_i \Theta_i d_i + X' \pi_R + v_R \quad (8)$$

$$L = \sum_i \gamma_i d_i + X' \pi_L + v_L \quad (9)$$

where:

$$\Theta_i = (\partial_i + \beta_L \lambda_i) / (1 - \alpha_R \beta_L), \text{ and } \gamma_i = (\lambda_i + \alpha_R \partial_i) / (1 - \alpha_R \beta_L), i = 1, \dots, m.$$

Recalling that $\alpha_R < 0$ and $\beta_L > 0$, it is possible to conclude that $(1 - \alpha_R \beta_L) > 0$. If the denominators of the expressions for Θ_i and γ_i are both positive, then the signs of these terms depend on their numerators.

Although the structural equations are not identified, estimates of the reduced form equations can be used to provide information on the likelihood of differential treatment discrimination. Suppose that lenders raise the probability of rejection for applicants who are members in the i th group, i.e. $\partial_i > 0$. If $\lambda_i = 0$, as is assumed by current statistical approaches to testing for discrimination, reduced form estimates of Θ_i will be positive while estimates of γ_i will be negative. In general, if $\lambda_i = 0$, differential treatment discrimination is indicated when Θ_i and γ_i have opposite signs. Conversely, if there is discrimination and Θ_i and γ_i have identical signs, then λ_i cannot be zero. But in this case, one cannot claim that a non-zero Θ_i indicates discrimination on the part of banks, for, if $\lambda_i \neq 0$, we get $\Theta_i \neq 0$ even when $\partial_i = 0$. Put another way, if Θ_i and γ_i have opposite signs, one cannot conclude that there is discrimination against group i .

The possibility of differential treatment by lenders can now be tested using this statistical approach and the data collected by the Boston FED. Table 1 contains estimates of reduced form equations (6) and (7). Following the practice of other investigators, problem observations were eliminated from

Table 1

| Variable | OLS Coefficient | (t statistic) | Logit | (t statistic) |
|----------------------------------|-----------------|---------------|---------|---------------|
| Constant | 0.67 | (44.86) | -4.06 | (-14.77) |
| RISK OF DEFAULT | | | | |
| Net wealth | -0.00001 | (-3.84) | 0.00003 | (0.47) |
| Consumer credit history | 0.0009 | (0.36) | 0.28 | (7.91) |
| Mortgage credit history | 0.05 | (7.10) | 0.26 | (2.15) |
| Public record history | 0.05 | (3.13) | 1.33 | (7.27) |
| Unemployment in region | -0.002 | (-1.17) | 0.08 | (2.60) |
| Self-employed | -0.01 | (-1.00) | 0.47 | (2.46) |
| PERSONAL CHARACTERISTICS | | | | |
| Married minority (MM) | 0.06 | (4.98) | 1.12 | (5.85) |
| Single minority male (SMM) | 0.05 | (2.82) | 1.35 | (5.19) |
| Single minority female (SMF) | 0.05 | (2.80) | 0.77 | (2.79) |
| Single non-minority male (SNM) | 0.02 | (1.46) | 0.57 | (3.11) |
| Single non-minority female (SNF) | -0.02 | (-1.93) | 0.21 | (0.82) |
| Separated minority (SM) | -0.03 | (-0.87) | -0.65 | (-1.03) |
| Separated non-minority (SN) | -0.04 | (-1.56) | -0.31 | (-0.62) |
| | R ² | 0.06 | | |
| Number of observations | 2,497 | 2,497 | | |

the data, particularly applications for multifamily and investor properties. The exogenous variables are net wealth, consumer credit history, mortgage credit history, public credit history, expected unemployment rate and self-employment.

Applicant characteristics are indicated by dummy variables that divide the applicants into distinct demographic groups that might either be expected to be the object of differential treatment by loan officers or to have different tradeoffs between LTV and expected probability of rejection. It is crucial that the vector of demographic dummy variables be sufficient to account for both differential treatment by lenders and differential ability to avoid rejection on the part of applicants. In this case, the applicant groups are: single minority female (SMF), single minority males (SMM), separated minorities (SM), married minorities (MM), single non-minority females (SNF), single non-minorities males (SNM), separated non-minorities (SN) and married non-minorities (MN). Separated applicants have special legal reasons to limit disclosure of information on their income and wealth.

The results in Table 1 show that Θ_i, γ_i pairs are certainly not opposite in sign as would be expected if differential treatment discrimination played an economically significant role in lending decisions. Quite the contrary, the results are consistent with a finding of no differential treatment or a finding that whatever the effects of differences in δ_i they are overwhelmed by differences in λ_i . Furthermore, note that virtually all dimensions of the applicant's demographic characteristics – race, gender and

marital status — are associated with significant differences in the reduced form rejection equation. This reflects the endogeneity of rejection in the lending process.²⁷ Applicants having greater ability to avoid rejection and/or greater distaste for rejection have lower rates of rejection. Overall it appears that being separated produces the lowest rejection rate. If this effect is interpreted as lender discrimination, it seems most unreasonable. However, recall that the separated have an incentive to conceal income and assets and hence to gain approval with the minimum of disclosure. Thought of in these terms, the partial effect of separation on rejection is perfectly understandable.²⁸

Although this discussion has concentrated on the rejection equation, similar points can be made about using single-equation APR and ex-ante default models to test for discrimination. The problem is that APR and the ex-ante probability of default, i.e. the probability of default at application, are jointly determined endogenous variables along with other loan terms. All the arguments about simultaneous equations bias and lack of identifying variables in the loan file suggest that estimated coefficients of minority variables in these equations also provide no reliable information on discrimination. Note that the testing performed in this section and illustrated in Table 1 is easily accomplished and could be done by lenders concerned with the possibility of differential treatment discrimination in their own loan portfolios.

It is intellectually satisfying to observe that arguments about the use of single-equation models based on data in loan files to test for discrimination ends with some degree of symmetry. Rejection and APR equations tend to produce false positive indications of discrimination for the same reason that default equations tend to produce false negative indications of discrimination. The point to take away from this discussion is that all three of these conventional approaches to testing for discrimination are unreliable because they are not based on economic theory. The first two conventional tests tend to produce misleading results in the form of false positive indications of discrimination and the third tends to produce false negative indications of equal treatment. Given this bias, conventional estimates of rejection, pricing and default equations may still have some utility. Failure to get a positive indication of discrimination by getting a non-significant minority coefficient estimate in a rejection or APR equation indicates the absence of discrimination. Conversely, failure to get a positive and significant coefficient in a default equation may indicate a problem of discrimination. However, in the final analysis, all of these three conventional approaches to testing for discrimination are biased tests for bias in mortgage lending.

IV.E Critique of Conventional Models of Mortgage Credit and Prepayment Risk

The previous section discussed current estimates of mortgage default use in tests for discrimination in mortgage markets. In this section, current models of the joint determination of mortgage default and prepayment are discussed. These models are different statistically than single-equation default models, as can be seen by comparing the single-default equation (3) with the joint-default and prepayment

system in (4). Their characteristics and uses are discussed in detail in Van Order (2008) who notes that they have become very important in pricing mortgages and mortgage-backed securities and that their failure played a role in the current problems of mortgage lenders and the GSEs. Although the problems in these models are analogous to those discussed above in that they lack support in economic theory and assume mortgage terms are not selected strategically, their importance to the financial system suggests that they be given special treatment here.

Some aspects of conventional models of mortgage default and prepayment are extremely sophisticated and based on solid economic theory. These are the aspects that concern the incentive to default and prepay based on market conditions after endorsement. As noted in the literature review in III.B., conventional estimates of default and prepayment over the duration of mortgages have produced some very important results that agree well with economic theory. For example, the value of the option to prepay is well understood based on the mortgage interest rate, current market rates, interest rate volatility and the presence of prepayment penalties. Empirical estimates of the two-equation system in (4) also indicate that the presence of loan amortization, house price appreciation and the variance in house price appreciation all lower the probability of default as expected based on economic theory. Furthermore, changing local economic conditions, such as rising unemployment, tend to raise default. Put another way, variables reflecting changes in economic conditions after endorsement have theoretical effects on prepayment and default that have been modeled carefully using economic theory and estimates of current joint default, and prepayment models tend to confirm that these time-varying factors have the expected effects on observed default and prepayment.

Factors such as the movement of interest rates, house prices and local economic conditions are clearly independent of the default and prepayment decisions on individual mortgages. There is no basis in economic theory for believing that the household's probability of prepaying or defaulting on its own mortgage could influence either the national or local economy. In terms of the model in (4), this means that computation of the time-varying value of the options to default and prepay, B_{Dt} , and B_{Pt} respectively, can proceed using time-varying variables that appear to be exogenous to the decision to default. The time-varying economic conditions variables, E_t , also appear exogenous along with demographic characteristics. These time-varying covariates have received most attention in the academic literature and have been widely applied to mortgage markets ranging from automated underwriting, through pricing, to all aspects of secondary market operations as discussed in Van Order (2008).

The problem with current models of default and prepayment risk based on equation (4) is that the assumed exogeneity of the initial loan terms is in conflict with economic theory. Barth, Cordes and Yezer (1981) modeled default, LTV, term-to-maturity and payment-to-income (PTI) as jointly determined variables. Theory suggests that applicants who believe the probability of default is higher will tend to choose higher LTV, and longer term-to-maturity. They find, using a simultaneous equations model, that both of these theoretical expectations hold in the data. Even more important for the purposes here, they find that the estimated effects of LTV and PTI on default are very different in single-equation

models where these variables are assumed exogenous and in simultaneous-equation models of default. Overall, this suggests empirical support for the hypothesis that expectations for default influence applicants' choices of loan terms and implies that estimates of systems like (4) will be biased because they assume loan terms are exogenous.

These points have been recognized in the literature on modeling and evaluating credit and prepayment risk, but because there is no convenient way to modify conventional models they tend to be ignored. For example, Van Order (2008) has one paragraph on model validation which states:

“The increasing awareness of the moral hazard element of credit risk adds to the dimensions of validation problems. It is not enough to have models that fit past data well and make sense. The models (or modelers) also have to take into account the fact that if the models are to be used to evaluate credit risk and make lending or pricing decisions, people on the side of the deal will reverse-engineer the models and take advantage of omitted variables.” (pg. 20)

This warning concerns one manifestation of the general point about joint determination made here. In this case, the terms of the loan are engineered to lower the expectation of default produced by the model.

As noted above, Brueckner (1994a) considered the simplified case in which equally qualified applicants differed in default probability due to differential default cost in ways that were not observable by lenders. All aspects of the loan transaction were assumed equal except for LTV and APR. Even with these simplifying assumptions, theoretical analysis of the market proved far from trivial. Nevertheless, Brueckner was able to demonstrate that under reasonable conditions, lenders would not serve these two borrower types with the same contract in what is termed a pooling equilibrium. Instead, profit-maximizing lenders would offer loans in which LTV and APR vary directly. The high-default borrowers would be segregated into contracts with higher LTV and higher APR. This reinforces the Barth, Cordes and Yezer (1981) result in which default is endogenous because differences in default cost are causing differences in LTV and APR rather than loan terms causing default. In the earlier discussion of bias in estimates of a default equation used to test for discrimination, the emphasis was on the possibility that demographic characteristics of borrowers could be spuriously correlated with the unobservable default cost and bias estimates of the relation between race and default. Here, the concern is that estimates of the relation between LTV, APR, PTI, etc. and default in the model in (4) may be biased.

What is most troubling in light of recent experience with elevated levels of default and foreclosure, is the possibility that the ability of current statistical models to predict losses could have been compromised by the endogeneity of loan terms. Evidence on this is still accumulating but the structure of the argument is clear. One view is that current model performance has been satisfactory and that the rise in default on mortgages with given loan terms is due to changes in time-varying covariates, particularly falling house prices and deteriorating economic conditions.²⁹ Certainly falling prices put the default option

reflected in BD_t in the money and bad realizations of the economic, E_t , variables also raise default. The alternative view is that, when home prices are expected to be flat or falling and the expected probability of default rises, applicants modify their choices of loan terms, taking into account the possibility of future default. This type of modification may take three forms: reverse engineering of the credit risk or scoring equations discussed in Van Order (2008); fraudulent behavior of the type identified by Carrillo (2009) in which applicants with no intention of making mortgage payments choose the highest LTV available regardless of APR; or general shift in the relation between LTV and APR for all applicants as they now perceive that the probability of default is higher or more important in their choice of mortgage terms. Recently, Rajan, Seru and Vig (2010) have made an equivalent argument for endogeneity of loan terms by constructing a model in which “soft” information used in the underwriting process can be manipulated by applicants, producing biased predictions from default models.

The rise in fraudulent behavior is particularly noteworthy because it is inconsistent with the theory behind the conventional default equation which assumes that high LTV causes borrowers to default on mortgage payments. In contrast, models with endogenous loan terms predict that borrowers intending to default will seek high LTV in order to maximize return from fraud. The Federal Bureau of Investigation reports that 1,571 fraud cases were opened in 2009 compared to 146 in 2004 and that total mortgage suspicious activity reports equaled 67,190 mortgages with more than \$1.5 billion in losses.

There is an easy test to determine which view of model performance is correct. It is possible to estimate models of default and prepayment using conventional techniques, assuming loan terms are exogenous, and using data from cohorts of mortgages endorsed in different years. Obviously default rates in more recent year cohorts will be much higher. If this rise in default rates is explained by the time-varying covariates, particularly home prices, then the first view of satisfactory model performance is correct. What will be termed the baseline hazard function will remain relatively constant. Alternatively, if the baseline hazard function rises significantly as cohorts approach the period when defaults rose, this indicates that model performance was problematic and that one or more of the three manifestations of endogenous loan terms had a significant role in the failure to anticipate loan losses. This question is sufficiently compelling that a number of research papers on the topic will eventually be done. At this point, the literature is incomplete. However, one interesting preliminary effort by Quercia and Tian (2010) indicates that shifts in the baseline hazard function were important in explaining the rise in defaults.

V. Conclusions

A common theme underlies this essay. The mortgage lending transaction is extremely complex and involves many dimensions. Applicants, loan officers, underwriters and secondary market participants make decisions based on payment to income ratios, LTV, cosigners, interest rates, points, prepayment penalties and the probability of default and prepayment. Economic theory holds that choices regarding any one of these variables plus the rejection decision, arise jointly so that all aspects of the mortgage transaction are determined simultaneously. Thus the outcome of a mortgage transaction involves the simultaneous consideration of many factors about which both the applicant and the lender must come to some mutual agreement. The complexity of the transaction is sufficiently great that current theoretical models of the mortgage transaction only consider two variables at a time and generally conclude that markets involve substantial self selection in order to deal with asymmetric information and moral hazard.

The complexity of the mortgage transaction has long been recognized in the empirical literature on modeling loan rejection, pricing and default. As noted in the discussion above, Barth, Cordes and Yezer (1981) and Maddala and Trost (1982) treated loan terms, loan amount, interest rate, monthly payment to income ratio, cosigners, prepayment penalties, default, prepayment, etc. as jointly determined endogenous variables.³⁰ Furthermore, there is also a substantial literature on the demand for mortgage debt arguing that mortgage amount and quantity of housing purchased are jointly determined. According to this literature, many households are “down payment constrained” in that their desired housing consumption exceeds the amount of housing for which they qualify given their ability to make a down payment. Certainly for such households, loan amount and value are jointly determined endogenous variables. Households that are down payment constrained are likely to have mortgage outcomes that are very different than those who are not constrained. Thus far the theoretical literature on the demand for mortgage debt has not begun to consider the effects of down payment constraints on mortgage terms.³¹

Current models of discrimination generally involve a single equation and fail to deal with the problem of joint causation of all loan terms including the expectation of future default. Current models of credit and prepayment risk deal elegantly with the effects of exogenous time-varying covariates but again

assume that loan terms are exogenous. The result of this failure to consider joint determination of loan terms and expected default is bias in tests for discrimination and potentially serious problems for estimation of credit risk.

This is not an unusual problem in empirical and applied research in economics. Getting the con out of economics has been recognized as a serious problem by the economics profession and in the academic literature for over 25 years. The recent Symposium in the *Journal of Economic Perspectives* attests to the continued presence of the problem in other areas of economic research. The serious limitations of current statistical approaches to testing for discrimination and credit risk in mortgage lending have likely contributed to recent problems in mortgage markets. If these limitations are not recognized and naïve reliance on them continues, current problems are likely to recur in the future. There are major gains to be made if the con can be removed from economic analysis of mortgage market discrimination and mortgage credit risk.

End Notes

1. This should be a major concern of both the buyer's agent and the seller's agent and indeed the seller should consider the possibility that the loan application fails before accepting the contract.
2. Leamer's paper is one of many in the 1980's that complain about the use of econometric models lacking empirical support and the failure to perform extensive robustness testing in order to make sure that results are not sensitive to the variables in the model or the functional form chosen for the model.
3. Ehrlich argued further that enforcement efforts could also have an effect on the murder rate but that the murder rate could cause greater enforcement effort. In short, the murder rate, capital punishment and enforcement efforts in an area were jointly determined variables. In the discussion here, the issue of enforcement effort is left out for simplicity.
4. Specifically, the authors note that "Terms, however, are interdependent , for down payment percentage and interest rate are likely to vary inversely. Consequently, terms of the loan should be estimated simultaneously to test for possible discrimination." Pg. 187.
5. See, for example the discussion in Avery, Beeson and Sniderman (1993).
6. See Munnell, Browne, McEneaney and Tootell (1992).
7. While the empirical approach is identical to the earlier work, the expanded dataset allows measurement of many more variables, including 38 variables not in HMDA data.
8. The premium is 300 basis points for a first lien and 500 basis points for a second lien secured by a home mortgage. Regulation C requirements were amended in 2002. The 2004 HMDA data, released officially in September 2005, were the first to include information on these "higher priced loans." A number of other changes were associated with the 2002 Regulation C changes, including expanded coverage of non-depository institutions, distinguishing lien status, structure type and preapproval. All these invite wider use of HMDA data in both rejection and mortgage pricing equations.
9. For a similar discussion of the argument of equation (2) for the case of subprime lending see Crews Cutts and Van Order (2004).
10. To the extent that overhead expenses and servicing costs, holding other factors constant, do not vary with loan amount, APR should fall at a decreasing rate with loan amount and this quadratic effect is observed in the empirical literature.
11. For example see the Consumer Federation of America (2005, 2006) for examples of this type of statistical analysis using only HMDA data.
12. Note that the test for discrimination is based on the marginal player not differences in the average African-American and white players on the team.

13. Mortgage securitization reflects the legitimate gains from diversifying away the unique risk associated with mortgage credit and ultimately makes more credit available at lower cost to borrowers. For fixed rate instruments, securitization is also important in allowing interest rate risk diversification.
14. See, Nichols, Pennington-Cross and Yezer (2007) for a discussion of the business model and explanation of some of the necessary conditions for it to work. In particular, they explain why subprime loans have both higher interest rates and higher rejection rates than prime loans.
15. Automation of the underwriting decision has facilitated electronic application in which the underwriting decision can be made without the loan officer or underwriter ever meeting the applicant. Under these circumstances, the applicant can reveal, conceal or declare whatever racial or ethnic characteristics seem most advantageous to the transaction.
16. For subprime mortgages, borrowers whose creditworthiness improves after endorsement have a strong incentive to refinance and hence repay. This leaves those whose creditworthiness has not improved in the pool and raises the likelihood of default compared to non-default terminations in the surviving loans.
17. See Kau and Keenan (1995), and references therein.
18. See, for example, Clapp, Goldberg, Harding and LaCour-Little (2001), Clapp, Deng and An (2006).
19. Regarding this point, please refer to the discussion of the problem of the con in economics in section II and particularly to the recent papers on this issue in the *Journal of Economic Perspectives*.
20. One might think that rejection is well observed but the difference among a rejected loan, an incomplete loan, and a withdrawn loan can be difficult to determine.
21. Obviously, if lenders were primarily concerned with avoiding default, virtually all applications would be rejected.
22. In addition to the lack of an experimental design and control variables, data on mortgage transactions are recorded by loan officers and underwriters for reasons completely unrelated to scientific research. In particular, loan officer compensation is based on loan production not on the quality or completeness of the data collected.
23. For a discussion of the theory of prepayment risk and points see Brueckner (1994b).
24. The mass points of mortgages at critical LTV values indicate this strategic choice by applicants. For example, and 80 percent LTV is often needed to avoid mortgage insurance and the distribution of LTVs has a mass point at 0.80.
25. Curiously, the 1992 Boston FED working paper had the extended discussion of coaching and referenced the Barth, Cordes and Yezer (1981) which argues against single-equation models of rejection and suggests that rejection, loan amount, LTV, monthly payment-to-income ratio, cosigners, etc. are all jointly dependent. In the version of the Boston FED working paper that appeared subsequently in the *American Economic Review*, references to coaching were dropped as was the reference to Barth, Cordes and Yezer (1981). The authors do claim, in a footnote, to have used instrumental variables to deal with the possibility that the LTV is endogenous. This implies a very strange reading of the simultaneity argument which holds that all variables under the control of the applicant are endogenous. Consider, for example, that variables included in the model as independent or exogenous include housing payment and debt burden ratios. But these variables are a function of loan amount. If loan amount and LTV are endogenous, then it is a mathematical certainty that all other independent variables that are a function of loan amount are endogenous. Therefore, it is logically inconsistent for the authors to “test” for the endogeneity of LTV and at the same time pretend that variables which are a function of loan amount, like payment and debt burden ratios are exogenous. It is even more curious, that Ladd (1998), in commenting on the Boston FED study and subsequent papers pointing out the problem of multiple endogenous variables, fails to note the logical contradiction in the test for simultaneous equation bias.
26. Furthermore, recall that other endogenous variables such as payment to income ratios, have simply been left out of the analysis because they are also jointly determined endogenous variables.

27. The variation in estimated coefficients could only be due to variation in the ∂I if there was differential treatment discrimination practiced at a different level against each demographic subgroup.

28. Note also the low coefficient of determination in the loan-to-value equation illustrates yet another problem for structural estimation. Even if the equation system were identified, lack of good instruments for LTV would make structural estimation results problematic.

29. Clearly the changes in composition of loan products influenced the rise of default. In particular, loans with limited documentation will have their own selection effect and attract high-risk borrowers to the home purchase and refinance markets.

30. For a more recent discussion of the implications of these early papers for estimation of models of discrimination and credit risk, see LaCour-Little (2001) and Zhang (2010).

31. Regarding the joint determination of demand for housing and mortgage debt, see Jones (1993), Follain and Dunsky (1997), Brueckner (1995) and LaCour-Little (2001).

References

- Angrist, Joshua D. and Jorn-Steffen Pischke. 2010. The credibility revolution in empirical economics: how better research design is taking the con out of econometrics. 24 (2), pp. 3–30.
- Ambrose, Brent W., Richard J. Buttimer Jr., and Charles A. Capone, 1997. Pricing Mortgage Default and Foreclosure Delay, *Journal of Money, Credit, and Banking*, 29 (3), pp. 314-325.
- Avery, Robert, Patricia E. Beeson, and Mark S. Sniderman, 1993. Accounting for racial differences in housing credit markets, Federal Reserve Bank of Cleveland Working Paper 9310.
- Avery, Robert B., and Glenn Canner, 2005. New information reported under HMDA and its application in fair lending enforcement. *Federal Reserve Bulletin*, pp. 344-394.
- Barth, James R., Joseph J. Cordes, and Anthony M. Yezer, 1980. Financial institution regulations, redlining, and mortgage markets, in *The Regulation of Financial Institutions*, Conference Volume 21, Federal Reserve Bank of Boston, (April 1980), pp. 101-143.
- Berkovec, James A., 1998. Discrimination, competition, and loan performance in FHA mortgage lending, *Review of Economics and Statistics*, 80 (2), pp. 241–250.
- Berkovec, James A, Glenn B. Canner, Stuart A. Gabriel, and Timothy H. Hannan. 1994. Race, redlining, and residential home mortgage performance. *Journal of Real Estate Finance and Economics*, 9 (3), pp. 263–294.
- Black, Harold, Robert L. Schweitzer, and Lewis Mandell, 1978. Discrimination in mortgage lending, *American Economic Review*, 68 (2), pp. 186–191.
- Bostic, Raphael, 1996. The role of race in mortgage lending: Revisiting the Boston FED study, working paper, Board of Governors of the Federal Reserve.

- Brueckner, Jan K. 1994a. Mortgage default with asymmetric information, *Journal of Real Estate Finance and Economics* 20 (3), pp. 251–274.
- Brueckner, Jan K. 1994b. Borrower mobility, adverse selection, and mortgage points, *Journal of Financial Intermediation* 3, pp. 416–441.
- Brueckner, Jan K. 1995. The demand for mortgage debt: Some basic results, *Journal of Housing Economics*, 3 (4), pp. 251–262.
- Capozza, Dennis R. and Thomas A. Thomson, 2005. Optimal stopping and losses on subprime mortgages. *Journal of Real Estate Finance and Economics*, 30 (2), pp. 115–131.
- Carr, James H. and Isaae F. Megbolugbe. 1993. The Federal Reserve Bank of Boston study on mortgage lending revisited, *Journal of Housing Research*, 4 (2), 277–313.
- Carrillo, Paul, 2009. Fools or crooks: testing for fraud in the residential mortgage market. Working Paper, George Washington University.
- Clapp, John M., Gerson M. Goldberg, John P. Harding, and Michael Lacour-Little, 2001. Movers and shuckers: interdependent prepayment decisions. *Real Estate Economics*, 29 (3), pp. 411–450.
- Clapp, John M., Yongheng Deng, and Xudong An, 2006. Unobserved heterogeneity in models of competing mortgage termination risks. *Real Estate Economics*, 34 (2), pp. 243–273.
- Consumer Federation of America, 2005. Subprime locations: patterns of geographic disparity in subprime lending, unpublished report.
- Consumer Federation of America, 2006. Subprime locations: patterns of geographic disparity in subprime lending, unpublished report.
- Courchane, Marsha. 2007. The pricing of home mortgage loans to minority borrowers: How much of the APR difference can be explained, *Journal of Real Estate Research*, pp. 399–440.
- Crews Cutts, Amy, and Robert Van Order. 2004. On the economics of subprime lending, Freddie Mac Working Paper 04-01.
- Danis, Michelle A., and Anthony Pennington-Cross, 2008. The delinquency of subprime mortgages, *Journal of Economics and Business*, 60 (1–2), pp. 67–90.
- Day, Theodore E. and Stanley J. Liebowitz. 1998. Mortgage lending to minorities: Where’s the bias?, *Economic Inquiry*, 36 (1), pp. 3–28.
- Deng, Yongheng, John M. Quigley, and Robert Van Order, 2000. Mortgage terminations, heterogeneity, and the exercise of mortgage options. *Econometrica*, 68 (2), pp. 275–307.

Ehrlich, Isaac. 1975. The deterrent effect of capital punishment: a question of life and death. *American Economic Review*, 65 (3), pp. 397–417.

Follain, James R. and Robert M. Dunsky. 1997. The demand for mortgage debt and the income tax, *Journal of Housing Research*, 8 (2), pp. 155–159.

von Furstenberg, George M., 1969. Default Risk on FHA-insured home mortgages as a function of the terms of financing: a quantitative analysis. *Journal of Finance*, 24 (3), pp. 459–477.

Glennon, Dennis and Mitchell Stengel, 1994. An evaluation of the Federal Reserve Bank of Boston's study of racial discrimination in mortgage lending, Economic and Policy Analysis Working Paper 95-1, Comptroller of the Currency, Washington, D.C.

Haavelmo, Trygve. 1944. The probability approach in econometrics, *Econometrica*, 12, 1–115.

Harrison, Glenn W., 1998. Mortgage lending in Boston: a reconsideration of the evidence. *Economic Inquiry*, 36 (1), pp. 29–38

Ho, Giang and Anthony N. Pennington-Cross. 2006. The impact of local predatory lending laws on the flow of subprime credit, *Journal of Urban Economics* 60(2), pp. 210–228.

Ho, Giang and Anthony Pennington-Cross. 2006b. The termination of subprime hybrid and fixed rate mortgages. Working Paper 2006-042A St. Louis, Missouri: Federal Reserve Bank of St. Louis.

Horne, David K. 1997. Mortgage lending, race, and model specification, *Journal of Financial Services Research*, 11, pp. 43–68.

Hunter, William C. and Mary Beth Walker. 1996. The cultural gap hypothesis and mortgage lending decisions. *Journal of Real Estate Finance and Economics*, 13 (1), pp. 57–70.

Jones, Lawrence D. 1993. The demand for home mortgage debt, *Journal of Urban Economics*, 33 (1), pp. 10–28.

Kau, James B., and Donald C. Keenan. 1995. An overview of the option-theoretic pricing of mortgages. *Journal of Housing Research* 6, pp. 217–244.

LaCour-Little, Michael, 2001. A note on identification of discrimination in mortgage lending, *Real Estate Economics*, 29 (2), pp. 329–335

Ladd, Helen F. 1998. Evidence on discrimination in mortgage lending. *Journal of Economic Perspectives*, 12 (2), pp. 41–62.

Leamer, Edward, E. 1983. Let's take the con out of econometrics. *American Economic Review*. 73 (1), pp. 31–43.

- Leamer, Edward, E. 2010. Tantalus on the road to asymptopia. *Journal of Economic Perspectives*, 24 (2), pp. 31–46.
- Maddala, G.S., and Robert Trost, 1982. On measuring discrimination in loan markets, *Housing Finance Review*, 1, 245–268.
- Madden, Janice Fanning, 2004. Differences in success of NFL coaches by race, 1990–2002: Evidence of last hired first fired. *Journal of Sports Economics*, 5 (1), pp. 6–19.
- Mixon, Jr., Franklin G., and Len J. Trevino, 2004. How race affects dismissals of college football coaches, *Journal of Labor Research*, 25 (4), pp. 645–656.
- Munnell, Alicia H., Lynn E. Browne, James McEneaney, and Geoffrey M.B. Tootell, 1992. Mortgage lending in Boston: Interpreting the HMDA data, Federal Reserve Bank of Boston Working Paper No. 92-7.
- Munnell, Alicia H., Geoffrey M.B. Tootell, Lynn E. Browne, and James McEneaney, 1996. Mortgage Lending in Boston: Interpreting HMDA Data,” *American Economic Review*, pp. 25–53.
- Nichols, Joseph, Anthony Pennington-Cross and Anthony M. Yezer (2005). Borrower self-selection, underwriting costs, and subprime mortgage credit supply, *Journal of Real Estate Finance and Economics*, 30 (2), pp. 197–219.
- Pennington-Cross, Anthony. 2006. The duration of foreclosures in the subprime mortgage market: a competing risks model with mixing, Working Paper 2006-027A. St. Louis, Missouri: Federal Reserve Bank of St. Louis.
- Quercia, Roberto G., Michael A. Stegman, and Walter R. Davis. 2005. The impact of predatory loan terms on subprime foreclosures: the special case of prepayment penalties and balloon payments. Working Paper, Center for Community Capitalism at the University of North Carolina at Chapel Hill.
- Rothschild, Michael, and Joseph E. Stiglitz, 1976. Equilibrium in competitive insurance markets: an essay in the economics of imperfect information, *Quarterly Journal of Economics*, 80, pp. 629–649.
- Rachlis, Mitchel B. and Anthony M. Yezer, 1993. Serious flaws in statistical tests for discrimination in mortgage lending, *Journal of Housing Research*, 4 (2), pp. 315–336.
- Rajan, Uday, Amit Seru, and Vikrant Vig. 2010. Statistical default models and incentives. *American Economic Review*. 100 (2), pp. 1–5.
- Rose, Morgan J. 2008. Predatory lending practices and subprime foreclosures — distinguishing impacts by loan category. *Journal of Economics and Business*, 60 (1), pp. 13–32.
- Van Order, Robert. 2008. Modeling and evaluating the credit risk of mortgage loans: a primer. *The Journal of Risk Model Validation*. 2 (2), pp. 1–20.

Yezer, Anthony M., Robert F. Phillips, and Robert P. Trost. 1994. Bias in estimates of discrimination and default in mortgage lending: the effects of simultaneity and self-selection. *Journal of Real Estate Finance and Economics*, 9(3), pp. 197–215.

Zhang, Yan. 2010. Fair lending analysis of mortgage pricing: Does underwriting matter? Economics Working Paper 2010-1, Office of the Comptroller of the Currency.

ANTHONY YEZER

Professor Anthony Yezer is a member of the Department of Economics of The George Washington University where he directs the Center for Economic Research. He teaches courses in regional economics, urban economics, money and banking, the economics of crime and microeconomic theory. Three of his former students have won the AREUEA dissertation prize competition. His research interests have included the measurement and determinants of credit risk in lending, the effects of regulations on credit supply, migration, optimal city size and urban sprawl. He has been a Fellow of the Homer Hoyt School of Advanced Studies in Real Estate and Urban Economics since 1991.

Professor Yezer has been a consultant to the World Bank, various government agencies, and mortgage lenders. In his capacity as a consultant to government, he served as an expert witness for the Federal Trade Commission testifying in connection with trade regulation. He devised the current scheme for spatial pricing of postal services and regularly advises the U.S. Postal Service on updating their charges for post office boxes. He recently served on the National Research Council's Panel on Disaster Research Needs in the Social Sciences.

His articles have appeared in the *Journal of Finance*, *Journal of Real Estate Finance and Economics*, *Economic Inquiry*, *Journal of Law and Economics*, *Real Estate Research*, *Journal of Economic Perspectives*, *Land Economics*, *Journal of Economics and Business*, *Journal of Urban Economics*, *Journal of Regional Science*, *Regional Science and Urban Economics*, and *Southern Economic Journal*. He edited the monograph *Fair Lending Analysis: A Compendium of Essays on the Use of Statistics*, currently serves on the editorial boards of five journals and is editor of the American Real Estate and Urban Economics Association monograph series.

Professor Yezer received a Bachelor's degree from Dartmouth College, continued his studies at the London School of Economics and Political Science where he earned an M.Sc. degree and holds a Doctoral degree from the Massachusetts Institute of Technology. He was a Rhodes Scholarship finalist and received a National Collegiate Athletic Association Scholar-Athlete Fellowship.



1717 Rhode Island Ave., NW, Suite 400
Washington, DC 20036
www.mortgagebankers.org