THE PRICING OF CREDIT RISK IN PERSONAL AND CORPORATE LOAN MARKETS:
A THEORETICAL AND EMPIRICAL EVALUATION

by

WILLIAM ERIC MORAN
B.A., Wesleyan University (1981)

Submitted to the Department of Economics in Partial Fulfillment of the Requirements of the Degree of

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 1988

© William Eric Moran 1988

The author hereby grants to M.I.T. permission to reproduce and distribute copies of this thesis document in whole or in part.

Signature of Author

Department of Economics
February 17, 1984

Certified by

Julio J. Rotemberg
Thesis Supervisor

Accepted by

Peter Temin, Chairman
Departmental Graduate Committee

ARCHIVES
THE PRICING OF CREDIT RISK IN PERSONAL AND CORPORATE LOAN MARKETS:
A THEORETICAL AND EMPIRICAL EVALUATION

by

WILLIAM ERIC MORAN

Submitted to the Department of Economics
on February 17, 1988 in partial fulfillment
of the requirements for the Degree of
Doctor of Philosophy in Economics

ABSTRACT

The first essay examines the degree to which lenders price discriminate on the basis of observable personal characteristics in personal loan markets. This essay diverges from the mainstream literature by calculating loan rate differentials instead of attempting to estimate the degree of credit rationing in the loan market. The equilibrium loan rate is derived as an explicit function of the personal characteristics of borrowers and this equation is then estimated using a reduced form approach.

The hypothesis of no price discrimination is rejected at all standard significance levels. The individual characteristics that have the greatest influence on the interest rate are (a) one's race and (b) whether one has had a bill turned over to a collection agency or not.

The second essay tests (a) whether the simple life-cycle model of asset holdings holds and (b) whether credit constraints have an effect on asset holdings over the life-cycle. Under most plausible assumptions the time path of asset holdings should be hump-shaped, but the empirical literature has not confirmed this proposition. Also, credit constraints should lower an individual's asset holdings at any point in the life-cycle. These propositions are tested by regressing the net wealth to permanent income ratio on age growth.
variables and a variety of individual-specific variables.

The time path of the wealth to permanent income ratio is found to be hump-shaped, but not in a statistically significant manner. Conversely, the credit constraint variables are found to have a strong and statistically significant effect on the wealth to permanent income ratio. These results give support to the notion that individuals make decisions according to the life-cycle theory and are often constrained in doing so.

The third essay tests whether the corporate bond market efficiently incorporates all information about firm prospects in setting bond prices. More precisely, the essay tests whether the bond market uses all information in predicting corporate loan defaults. If the market is efficient then no information apart from the yield spread between an appropriately chosen Treasury bond and the corporate bond should have marginal predictive content in a bond default prediction model.

Efficiency is tested using a multivariate logit model employing the yield spread, other bond market related variables and accounting and stock market variables. Efficiency is rejected using a variety of methods. The most important variable except for the yield spread in predicting loan defaults is firm dividends as a percentage of total assets. These results are robust to (a) alterations in the lag between the release dates of the financial information and the default dates (b) different sample periods (c) the use of exchange versus non-exchange prices and (d) different valuations of bond covenants.

Thesis Supervisor: Julio Rotemberg, Professor of Applied Economics, Sloan School of Management
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS ............................................................... 5

CHAPTER ONE. INDIVIDUAL CHARACTERISTICS AND THE PRICING OF CREDIT RISK IN PERSONAL LOAN MARKETS ....... 6
  Introduction
  Consumer Maximization Model with Default Possibilities
  Loan Market Equilibrium
  Sources of Price Discrimination in Loan Market
  Discussion of Test Design
  Conclusion
  Appendix
  Footnotes
  References

CHAPTER TWO. CREDIT CONSTRAINTS AND THE TIME PATH OF ASSET HOLDINGS OVER THE LIFECYCLE ....................... 55
  Introduction
  Constructing a Model of the Wealth-Age Relationship
  Altering the Model to Account for Credit Constraints
  Estimates of the Wealth-Age Relation
  Footnotes
  References

CHAPTER THREE. CORPORATE LOAN DEFAULT AND THE PRICING OF RISK IN THE CORPORATE BOND MARKET ............... 96
  Introduction
  Bond Market Efficiency Model Developed
  Timing Issues Discussed
  Construction of Bond Yield Spread Model
  Multivariate Logit Test Design Developed
  Rejection of Bond Market Market Efficiency
  Robustness of Rejection of Market Efficiency
  Conclusion
  Footnotes
  References

BIOGRAPHICAL NOTE ............................................................. 139
ACKNOWLEDGEMENTS

I am indebted to a number of people for help with this dissertation. Julio Rotemberg and Bob Pindyck have helped me with incisive suggestions and helpful comments at every stage of this dissertation. Julio provided me with theoretical insights and suggestions that greatly improved this dissertation. Bob taught me a great deal about empirical research when I was his research assistant and his lessons were valuable in improving the empirical content of this dissertation. Without their help, writing this dissertation would have been much more difficult. I would also like to thank Jim Poterba for quickly and efficiently reading the dissertation as the third reader.

I would also like to thank Jeff Rosensweig. I greatly appreciated his friendship and support throughout my academic career at MIT. I have enjoyed both his keen mind and amusing manner.

I would also like to thank the numerous students I have taught at both Sloan and the undergraduate level during my stay at MIT. They were a constant source of inspiration for me between 1983 and 1985. I enjoyed both their friendship and intelligence during this period.

Fellowships and teaching positions I have received from MIT have made completing this dissertation an easier task. I am grateful to MIT for its financial support.
INDIVIDUAL CHARACTERISTICS AND THE PRICING OF CREDIT RISK IN CONSUMER LOAN MARKETS
I. Introduction

The purpose of this paper is to determine the effects of individual characteristics on the pricing of consumer credit. Most empirical credit market research has focused on credit rationing but personal loan markets are typically characterized by both credit rationing and price discrimination. Determining the nature and extent of price discrimination is important because (1) discrimination on the basis of certain individual characteristics such as race is not legal and (2) appropriate public policies aimed at the personal loan market are difficult to formulate without such knowledge.

Remarkably little direct evidence has been produced on the degree to which lenders price credit risk in personal loan markets. Most empirical credit market research has centered on the extent to which credit is rationed according to borrower credit risk levels. Paxson (1985) finds that individual borrowing levels are strongly affected by interest rate ceilings and labor income. Barth et al. (1983) conclude that state-specific legal variables (such as the maximum amount of income
that is garnishable in the case of personal bankruptcy) have a significant effect on personal loan rates, but they do not explicitly consider the effects of personal characteristics on loan rates.

While there is a considerable body of literature on credit rationing, there is very little empirical research which addresses how lenders both ration credit and price discriminate as a response to borrower riskiness. It would be optimal to simultaneously test an empirical credit market model that accounted for both credit rationing and price discrimination since it is well-known that both of these mechanisms are used by lenders in personal credit markets. In theory both the amount loaned and the interest rate on the loan are affected by a potential borrower's default probability. Stiglitz and Weiss (1981) prove that it is optimal for lenders to ration credit and price discriminate when there are many borrowers, some of whom are observationally indistinguishable. Modigliani and Jaffee (1970) argue that market imperfections and transactions costs make it optimal for financial institutions to segment the lending market into a small number of segments and charge the
same interest rate within each segment.

The lack of research in this area is attributable in part to the fact that the empirical testing of a rationing/price discrimination model is very difficult. One reason for this difficulty is that it is often impossible to compile the necessary data on all borrowers and would-be borrowers. Testing the model would require information on interest rates on loans granted as well as personal characteristics on borrowers and would-be borrowers. Usually, however, data sets contain interest rates or loan rejection information but not both. A second problem is that consistent parameter estimates may be generated only if there are no unobserved variables determining which agents receive credit and which do not.\(^{(2)}\)

This paper diverges from the mainstream credit literature by explicitly testing and analyzing a price discrimination oriented model of the consumer loan market. A simple two period model of consumption and borrowing is developed in which an individual chooses his borrowing level with an uncertain second period income and where, under certain circumstances, he may find it
optimal to default in the second period. The supply side is modeled with a similarly simple scheme. Suppliers are assumed to be risk neutral and to operate in a competitive market. The market equilibria are then derived with relatively few restrictions on the actions of borrowers and lenders. The equilibrium loan rate is derived as an explicit function of the personal characteristics of borrowers and is then estimated using a reduced form approach. This model permits an explicit statistical test of (1) whether lenders charge significantly different interest rates to low and high wealth borrowers and (2) whether lenders price discriminate on the basis of borrower individual characteristics, such as race and marital status. Neither of these hypotheses have been rigorously tested in previous research.

The major results of this study can be summarized as follows. First, a borrower's race and whether he has had a bill turned over to a collection agency, have a strong and statistically significant impact on loan rates that he faces (if he is a low-wealth borrower). The statistical power of the collection agency variable
confirms the hypothesis that lenders are sensitive to risk variations among lower wealth individuals. The significance of the race variable is similar to the Avery (1982) finding that blacks are constrained from borrowing their desired levels in the loan market. Second, the hypothesis that high and low income individuals face identical interest rates is rejected at all standard significance levels. This result provides evidence that quantity or credit rationing models provide an incomplete picture of the loan market because in a pure rationing model there should be no price discrimination. Third, loan rates to high wealth individuals are not priced on the basis of individual characteristics and the interest rate is constant within the high wealth subsample. These results, must be interpreted with caution because (1) the model emphasizes interest rate differentials to the exclusion of borrowing level differentials and (2) reduced form rather than structural form equations are estimated.

The structure of this paper is as follows. A theoretical model of the loan market is presented in sections II, III and IV, loan market equilibrium is
discussed in section V, the regression model is analyzed in section VI, the econometric results are presented in section VII and the conclusion is presented in section VIII.

The Loan Market

II. The borrower's budget:

I first analyze the primary elements of the demand side of the loan market. I initially focus on the consumption and borrowing choices of a representative individual, covering his time of life, borrowing possibilities and income distribution.

Each individual lives for two periods, with a labor income of \( y_1 \) in the first period and \( y_2 \) in the second period. First period income is known with certainty when the individual chooses first period consumption \( c_1 \). Second period income \( y_2 \) is stochastic and is not known to the individual until the start of the second period.

The individual may purchase nondurable goods in periods 1 and 2, yielding consumption levels of \( c_1 \) and \( c_2 \) respectively. First period income may be lower than desired first period consumption \( c_1 \). If \( c_1 > y_1 \) then the individual may borrow amount \( B \) at interest rate \( r \) to make
up the difference between $c_1$ and $y_1$. The individual is obligated to repay $rB$ in the second period, but there is a possibility that the individual will default on his loan.

A borrower defaults if $y_2 - rB$, second period income minus the obligated repayment, is lower than some floor level of income protected by bankruptcy. For simplicity, I assume that the floor level is fixed at $y_{\text{min}}$ for all borrowers.

The possibility of default changes the second period budget constraint because the repayment of $rB$ is no longer guaranteed. First and second period budget constraints when default is a possibility are given by

$$c_1 = y_1 + B$$

(1)

$$c_2 = \max[y_{\text{min}}, y_2 - rB]$$

(1a)

where the right-hand term in the brackets is the individual's consumption level if he doesn't default. An individual will default if $U(y_{\text{min}}) > U(y_2 - rB)$, the utility from defaulting is greater than the utility from not defaulting.\(^{(3)}\) The individual therefore defaults when $y_2 < y_{\text{min}} + rB$.

Another critical element besides the budget
constraints that determines an individual's borrowing opportunities is his income distribution in the second period of his life (first period income is known with certainty). The density function of $y_2$ is given by $g(y_2)$. I assume that $y_2$ is distributed uniformly:

$$g(y_2) = \frac{1}{(b - a)} \quad a \leq y_2 \leq b \quad (2)$$

The mean and standard deviation of $y_2$ are

$$\bar{y} = \frac{(b + a)}{2}$$

$$\sigma = \frac{(b - a)}{\sqrt{12}} \quad (4)$$

It is assumed that $y_{\text{min}} > a$ for simplicity. The mean ($y_u$) and approximate range ($\sigma_u$) of income above the floor level $y_{\text{min}}$ are given by

$$y_u = \frac{1}{2}(b + y_{\text{min}}) \quad (5)$$

$$\sigma_u = \frac{(b - y_{\text{min}})}{\sqrt{12}} \quad (5a)$$

The distinction between $\sigma$ and $\sigma_u$ may be critical, even though income cannot technically fall below $y_{\text{min}}$. If $\sigma/\sigma_u$ is very large an individual will probably default and receive $y_{\text{min}}$ and if $\sigma_u/\sigma$ approaches zero an individual will not default.
III. Utility Maximization:

Each individual chooses B in order to maximize the sum of first period utility and expected second period utility. Individuals are assumed to borrow if and only if \( b > y_{\text{min}} + rB \), that is, if there is a positive probability of repayment. \(^{(3a)}\) If the maximum that an individual can earn is lower than the default income level, the individual will not borrow. \(^{(4)}\) The range of permissable solutions for B is thus given by

\[-y_{1} \leq B < \frac{(b - y_{\text{min}})}{r}\]  \hspace{1cm} (6)

The upper boundary is the maximum that an individual could borrow with a positive probability of repayment and the lower boundary is the maximum that an individual could save (negative borrowing).

The discount factor is \( \delta \) and the individual's period utility function \( U(\cdot) \) is assumed to be quadratic:

\[ U_i = d c_i - \frac{1}{2} c_i^2 \]  \hspace{1cm} (7)
where $U_i$ (i=1,2) represents total utility in period $i$ and $d > c_i$ for all $c_i$. The individual's maximization problem may be posed as

$$
\text{Max } \frac{d(y_1 + B) - 1/2(y_1 + B)^2 + 1/\delta \int_{\text{a}} [dy_{\text{min}} - 1/2y_{\text{min}}^2]g(y_2)dy_2}{B}
+ \int_{\text{a}} [dy_{\text{min}} - 1/2y_{\text{min}}^2]g(y_2)dy_2
+y_{\text{min}}
\int_{\text{b}} [d(y_2 - rB) - 1/2(y_2 - rB)^2]g(y_2)dy_2
\int_{\text{z}}
$$

where $z = y_{\text{min}} + rB$, the highest income level at which an individual will default.

Each individual is assumed to be a Cournot-Nash participant in the loan market. That is, he assumes that lenders will not alter their interest rates in response to a change in his desired borrowing level (i.e., $\partial r/\partial B = 0$). If individuals correctly assumed that the interest rate charged is a positive function of the amount borrowed, none of the qualitative results of the model would be altered but the mathematics would be substantially complicated.\(^5\)

Under the assumptions specified above, the first
order condition for a maximum is

\[ d - B - y_1 + \frac{1}{\sigma} \delta \{ \text{drymin} - r/2 \text{ymin}^2 - \text{drymin} - dr^2 B - \]

\[ dr \sigma_u + 2 dr^2 B + r \sigma_u y_u - r^2 \text{Bymin} - \frac{1}{2} r^3 B^2 - \sigma_u r^2 B + r^3 B^2 \]

\[ + \frac{r}{2 \text{ymin}^2} \} = 0 \]

(9)

Equation (9) is quadratic in \( B \) and the solution is given by

\[ B = \frac{1}{r^3} \left( \delta \sigma - r^2 (d - b) - \sqrt{\delta^2 \sigma^2 - 2 r^3 [\delta \sigma (d - y_1) - drb + ry_u \sigma_u]} \right) \]

(10)

The negative root is chosen because it is the root that ensures that the second order conditions for a maximum hold (the solution is a global maximum given the assumptions of the model). The negative root also makes economic sense since \( B \) is a positive function of expected second period resources and a negative function of first period resources. The root chosen is necessarily real because the boundary conditions for \( B \) are violated when
the roots are complex.\(^{(6)}\) The demand for loans is downward sloping at all points. It might seem that \(B\) could increase with \(r\) from inspection of (10), but this can occur only when the boundary or endpoint conditions are violated.

The interior solution for \(B\) may be simplified if we assume that \(d = b\):

\[
B = \frac{1}{r^3} \left( \delta \sigma - \sqrt{\delta^2 \sigma^2 - 2r^3 [\delta \sigma (b - y_1) - rb^2 + r \sigma_u y_u]} \right)
\]

(11)

The optimal borrowing level (which may be negative) is increasing in \(\sigma\), \(\sigma_u\), \(\delta\) and \(y_u\) and decreasing in \(y_1\) and \(r\). None of these results are at variance with modern credit market theory. For example, an increase in first period income loosens a borrower's first period budget constraint and thus his demand for loans decreases.

Since \(B\) may be negative (the individual saves) it is useful to determine the cutoff point which determines whether an individual borrows or not. The minimum level of \(\sigma\) which induces an individual to borrow is given by:

\[
\sigma^* = \frac{r}{\delta [y_{\text{min}}^2/(b - y_1)]}
\]

(11a)
\( \sigma^* \) is affected by the right-hand side variables in the expected manner. For example, if an individual has low first period resources relative to his maximum potential income level and a low income variance, then he is likely to borrow. Also, \( \sigma^* \) is increasing in \( r/\delta \): the higher the price of future consumption, the higher one's income variance needs to be before one borrows. If \( r/\delta \) is very large, an individual will borrow only if \( b - y_1 \) is much larger than \( y_{\text{min}}^2 \).

While the model developed above gives a pretty complete description of borrower actions, it still does not contain the adverse incentive effects that the Stiglitz and Weiss (1981) model possesses. In that model, borrowers have incentives to undertake riskier projects as the interest rate increases. In this model, a borrower's riskiness \( \sigma \) is presumed to be independent of his borrowing activities. It is not clear, however, that the absence of adverse selection effects will have a significant effect on the empirical implementation of the model developed above. The appendix contains an adverse selection model that carries similar empirical
interpretations as the model developed in this section.

IV. Supply Side of the Loan Market

I assume that there are a large number of lenders and borrowers in the loan market. I further assume that lenders can observe certain individual characteristics which allow them to charge different rates to individuals of different creditworthiness. Riskier borrowers therefore pay higher interest rates than safe borrowers.

Firms are risk neutral and thus attempt to maximize the expected returns from serving borrowers. It is assumed that each lender does not know how much an individual borrows at other lending institutions. I further assume that a lender does not incur costs in processing a loan. The opportunity cost to each lender is assumed to be $r$.

The per dollar return to a lender lending amount $B$ at interest rate $r$ to a borrower is

$$\pi = \min[(y_2 - y_{\text{min}})/B; r]$$

(12)

The borrower must pay back either $rB$, the promised
amount, or \( y_2 - y_{\text{min}} \). If the individual's income falls below \( y_{\text{min}} \) the lender receives nothing.

The expected return to a lender from lending to an individual is

\[
E(\pi) = \frac{b}{z} \int (y_2 - y_{\text{min}})/B g(y_2) dy_2 + r \int g(y_2) dy_2 \quad \text{(13)}
\]

where \( z = y_{\text{min}} + rB \). The first term in (13) represents the expected profit to the lender when the borrower defaults and the second term represents the expected return to the lender when the borrower repays the obligated amount.

Equation 13 may be rewritten as

\[
E(\pi) = \frac{r \sigma_u}{\sigma} - \frac{1}{2} r^2 B / \sigma \quad \text{(14)}
\]

Equation 14 is quadratic in \( r \). \( E(\pi) \) approaches \( r \) as \( B \) approaches zero only if \( \sigma_u / \sigma \) approaches one. Lending even an infinitesimally small amount to someone who is nearly broke involves risk, and this is reflected in the first term in (14).

The comparative static results are relatively straightforward. The \( \partial r / \partial \sigma \) and \( \partial r / \partial B \) terms are positive.
The effect of $\sigma$ on $r$ is given by

$$\frac{\partial r}{\partial \sigma} = \frac{1/2(r^2)\partial B/\partial \sigma + r\sigma_u - 1/2r^2B}{\sigma \partial E(\pi)/\partial r} > 0$$

(17)

Lender and borrower actions combine to generate this result—lenders penalize a higher income variance at a fixed borrowing level and individuals borrow more as $\sigma$ increases.

The sign of $\partial r/\partial b$ is ambiguous. For a fixed borrowing level a lender will lower the interest rate charged as $b$ increases. The borrower increases his optimal borrowing level as $b$ increases, however, leading to a higher interest rate. The net change in $r$ will be negative only if the borrower response to $b$ is weak.

The expected return to the lender is monotonically increasing in $r$. This result is relatively unusual in the credit market literature. Stiglitz and Weiss (1981) show that the expected return to a lender is typically rising at first and then falls after a critical interest
rate \( r^* \) is reached. In this model \( \partial E(\pi)/\partial r \) is generally positive:

\[
\frac{\partial E(\pi)}{\partial r} = \frac{b - y_{\min} - rB - 1/2(r^2)\partial B/\partial r}{\sigma} > 0 \tag{16}
\]

\( \partial E(\pi)/\partial r \) is positive because \( b - y_{\min} - rB \) is positive by the assumptions of the model and demand is downward sloping (\( \partial B/\partial r < 0 \)). Since \( E(\pi) \) is monotonically increasing in \( r \), the type of credit rationing described by Stiglitz and Weiss (1981) is precluded. A sufficient condition for rationing in this model (and theirs) is an \( E(\pi) \) with an interior maximum (i.e., \( E(\pi) \) is nonmonotonic in \( r \)).

Rationing is impossible if the \( E(\pi) \) function is increasing in \( r \) because the lender never has an incentive to lower the interest rate and ration credit. Figure 1 shows an \( E(\pi) \) function that is nonmonotonic in \( r \). If lenders initially charge an interest rate of \( r_i \), they can increase profits by lowering the rate to \( r^* \) and reducing
the amount lent at such an interest rate. If the $E(\pi)$ function is monotonically increasing in $r$, no such rationing equilibrium can occur because if the interest rate is above (below) $rr$, the opportunity cost of funds, competitive pressures will force the interest rate down (up) until it equals $rr$.

A rationing model could be constructed under slightly different assumptions about the loan market. I assume that each borrower's $\sigma$ is observable to the lender. If $\sigma$ is assumed to be unobservable to the lender, then rationing may occur (i.e. some agents will receive loans and some observationally equivalent agents will not). The difficulties involved in estimating such a rationing model are discussed in the appendix.

V. Loan Market Equilibrium

There are a variety of possible equilibria in a credit market with borrowers and lenders such as those described above. I assume that (1) all lenders have a fixed opportunity cost of funds given by $rr$, (2) there are a large number of both lenders and borrowers and (3) borrowers are heterogeneous. Under these assumptions equilibrium is
defined by the relation $E(\pi) = rr$. If a lender tried to charge a higher rate than $rr$, nobody would borrow from that lender and if he charged a lower rate than $rr$ he would have negative economic profits.

If there are heterogeneous borrowers, lenders price discriminate on the basis of observable personal characteristics of the borrowers. The key assumption underlying this result is that lenders are able to distinguish borrowers on the basis of such observable individual characteristics.\(^8\)

The formal representation of the equilibrium is

$$rr = r\sigma u / \sigma - 1/2 (r^2) B^* / \sigma$$  \hspace{1cm} (18)

where $rr$ is the opportunity cost of funds and $B^*$ is given by equation (11), the optimal borrowing level chosen by the borrower. Figure 2 shows equilibria in a loan market with three borrowers. There are three borrowers in the market, each paying a different interest rate. As expected, the interest rate paid is monotonically increasing in default risk; the borrower generating the lowest expected return to the lender at a given interest
rate (represented by $E(n^3)$) compensates by paying the highest interest rate $r^3$. This is not an entirely unreasonable result. Although risky borrowers will typically be credit rationed, they also typically pay higher rates on the credit that they are lent. (This holds true in the public corporate loan market as well, where risky borrowers pay higher rates than less risky borrowers and are sometimes rationed).

The equilibrium described by (18) presents a limited picture of the consumer loan market. The assumption that borrowers are allowed to borrow their desired levels is unrealistic for borrowers with low income levels. The main point of this model, however, is to determine how interest rates are affected by individual characteristics and it is not evident that rationing will qualitatively alter this relationship.⁽⁹⁾

VI. Regression Model

Estimating (18) does not allow us to disentangle the effects of demand and supply parameters on the interest rate charged. Although only reduced form parameters may be recovered by estimating (18), it is still possible to
make structural inferences if the exogenous variables affect the supply and demand sides in opposite manners. For example, if individuals who have had bills turned over to a collection agency have higher demands but face lower supplies of credit then the interest rate must be increasing in the collection agency variable.

Substituting equation 10 into 14 yields the equilibrium interest rate but not a closed form solution for $r$. A first order Taylor series approximation of (14) (after substituting (10) into (14)) generates the interest rate as a function of all of the relevant exogenous variables in the model:

$$ r = a_0 + a_1 \sigma + a_2 \delta + a_3 y_1 + a_4 y_u $$

(19)

$\partial r / \partial \delta > 0$, $\partial r / \partial y_1 < 0$ and $\partial r / \partial y_u > 0$, as calculated above. The formulation above does not allow us to disentangle the various channels through which a particular individual characteristic may affect the interest rate because some individual characteristics will affect the interest rate through more than one of the four variables listed above. If we assume that each of the
individual characteristics has a linear effect on all of the right hand side variables then equation (19) may be rewritten as

$$r = \alpha_0 + \sum_{i} \sum_{j} \alpha_i \beta_{ij} X_j$$

(20)

where $X_j$ is the vector of individual characteristic variables, $\beta_{ij}$ is the effect of the $j$th individual characteristic on the $i$th variable and $n$ is the number of individual characteristics in the $X$ vector.

To simplify the complexities involved in determining the effects of individual characteristics on the pricing of consumer credit, I consider two special cases of borrowers. The first type of borrower has a $\sigma_u/\sigma$ ratio that approaches one. Such an individual is considered a very good credit risk. For this type of borrower the partial derivatives of $r$ (derived from the equilibrium equation (18)) with respect to all of the variables approach zero. The interest rate should thus be constant across the high-wealth sample. A test of this
"constant interest rate" hypothesis may be conducted by examining the F test for all of the regression coefficients in equation (20).

The second type of borrower has a very low $\sigma_u/\sigma$ ratio. This type of individual is considered a very poor credit risk. His ability to repay a loan of any size is in doubt. As $\sigma_u$ approaches zero, the direct effects of the other variables on the interest rate charged to the second class of borrower become insignificant (i.e., the coefficients on the other variables in the interest rate equation become very small).

The considerations above suggest that low-wealth individuals face higher interest rates than high-wealth individuals. Also, we should expect the variance of interest rates faced by low income borrowers to be significantly higher than the variance of rates faced by high income borrowers (since interest rates should be relatively constant for high income individuals).

In summary, the primary hypotheses to be tested are (1) whether mean interest rates faced by high and low-wealth individuals are significantly different, (2) whether the variances of the rates faced by the high and
low wealth individuals are significantly different and (3) whether individual characteristics have statistically significant effects on the rates that individuals face. Although rate differentials do not prove the existence of price discrimination, they do provide evidence in its favor, especially if the individual characteristics that are statistically significant are variables related to creditworthiness.

Virtually none of the hypotheses listed above have been examined in any detail in previous research. Almost all credit market research has focused on credit rationing to the exclusion of price discrimination. The few studies that have analyzed interest rate determinants, (such as Barth et al. (1983)) have not analyzed how individual characteristics related to creditworthiness affect interest rates.

VII. Econometric Tests
a. Choice of Variables:

A wide variety of individual-specific and state-specific legal variables have been used to explain consumer borrowing levels (and to a lesser extent, loan rates) in past research. I will focus on the
individual-specific factors to the exclusion of the legal factors for a number of reasons. First, developing accurate measures of such variables as the loan rate ceiling is a difficult task. Second, legal measures have not performed well in previous regressions. Barth et al. (1983) found that most legal variables, such as state garnishment laws, had the wrong signs in reduced form interest rate regressions. Third, the potential specification error from excluding legal variables may be relatively low, given the low correlations between individual characteristics and legal variables (Barth et al. (1983)).

Another class of variables that is employed describes the structure of the consumer loan market. Variables such as the degree of concentration in a market and the distribution of types of lenders in a market (finance companies versus banks, for example) may explain a certain percentage of the variance in loan rates. Unfortunately, such data is extremely difficult to obtain. Also, the degree of competitiveness in a loan market may be a function of the individual characteristics of the borrowers in the market.
The variables used in this study are presented in Table 1. The first five variables (RESCH, COLLAG, BLACK, CHECK and TRES) are more supply than demand oriented. Previous studies have indicated that lenders tend to discriminate between borrowers mostly on the basis of these variables. Both Smith (1964) and Wiginton (1980) found the most important variables in explaining loan default were (in order of importance): housing status, bank account status, time spent in last residence and the purpose of the loan. Avery (1982) found race to be the critical factor in determining whether individuals were rationed in the loan market.

The next set of variables (URB, NEAST, NCENT and SOUTH) are weak proxies for local market structure. Peterson and Ginsburg (1981) found degree of urbanization to be important in explaining loan rates on autos. Loan rates in urban settings were typically lower than rates in rural settings. The reason suggested for this phenomenon is the greater degree of competition among financial institutions in urban settings. Further market influences may be captured by the geographic dummies. The remaining variables in
TABLE 1

The variables in the $y_u$ vector are:

RESCH: A dummy variable that takes on the value of one if the individual had to reschedule or renegotiate a debt in the previous year.
COLLAG: A dummy variable that takes the value of one if the individual has ever had a bill turned over to a collection agency.
BLACK: A dummy variable taking on the value of one if the individual is black.
CHECK: A dummy variable that takes on the value of one if the household has a checking account.
TRES: Time spent at current residence.
URB: An variable ranging from 1 to 6, where 6 represents a rural area and 1 represents one of the 12 largest metropolitan areas in the United States.
NEAST: A dummy variable that takes on the value of one if the household lives in the northeast.
NCENT: A dummy variable that takes on the value of one if the household lives in the north central region of the country.
SOUTH: A dummy variable that takes on the value of one if the household lives in the south.
LA: Liquid assets at the beginning of the period. The sum of amounts held in checking and savings accounts (and bonds).
INC: Disposable income
DATT: A measure of family attitudes towards holding debt. The higher the index, the more that the family considers borrowing "bad".
SEX: A dummy variable that takes on the value of one if the household head is female.
NP: Number of people in the household.
RENT: A dummy variable that takes on the value of one if the individual is currently renting an apartment or room.
AGE: Age of household head.
DIV: A dummy variable that takes on the value of one if the household head is divorced.
Table 1 may affect both borrower and lender behavior.

This study utilizes the "Consumer Durables and Installment Debt" ICPSR data set. This data set contains information on the amounts borrowed, interest rates and individual characteristics such as race, housing status and educational level for 1436 households over four years. All observations with complete data available were used, reducing the total sample to 1069 observations. The econometric results below may be biased if the "complete data" subsample has significantly different characteristics than the rest of the sample.

b. ESTIMATION

I divided the sample into high and low-wealth groups for all estimations. I split the sample by variables known at time t to avoid the selection bias that occurs when expectation variables in the loan rate equation are correlated with the splitting variable. The splitting variable is income plus liquid assets in the year preceding the current sample taken in 1970. The median value of income plus liquid assets is 12,130. To minimize the possibility of misclassification I remove
all observations where liquid assets plus income is between 8,000 and 16,000, approximately the middle third of the sample.

The means and standard deviations of the low and high income subsample interest rates are presented in Table 2. The mean interest rates faced by the low and high-wealth subsamples are significantly different from one another, 11.31 and 9.97 respectively. The null hypothesis that the means are equal may be rejected at the 1% significance level. This suggests that lenders practice price discrimination.

The two subsamples have significantly different interest rate variances, as predicted by theory. The low and high-wealth income subsample variances are 27.1 and 14.5 respectively and the null hypothesis that the difference is zero may be rejected at the 1% level. This provides evidence for the hypothesis that high-wealth individuals face relatively flat interest rate schedules compared to low-wealth individuals.

The reduced form estimations of equation (20) for the low and high-wealth sample are presented in Table 3. The results support the notion that lenders practice
### TABLE 2

<table>
<thead>
<tr>
<th></th>
<th>High Income Sample</th>
<th></th>
<th>Low Income Sample</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Dev.</td>
<td>Mean</td>
<td>Std. Dev.</td>
</tr>
<tr>
<td>INT. RATE</td>
<td>9.97</td>
<td>3.81</td>
<td></td>
<td>11.315</td>
</tr>
<tr>
<td>INC</td>
<td>20167.61</td>
<td>13051.19</td>
<td></td>
<td>6529.01</td>
</tr>
<tr>
<td>BLACK</td>
<td>.154E-01</td>
<td>.123</td>
<td></td>
<td>.201</td>
</tr>
<tr>
<td>COLLAG</td>
<td>.650E-01</td>
<td>.246</td>
<td></td>
<td>.142</td>
</tr>
<tr>
<td>URB</td>
<td>3.57</td>
<td>1.473</td>
<td></td>
<td>3.928</td>
</tr>
<tr>
<td>AGE</td>
<td>47.0</td>
<td>9.141</td>
<td></td>
<td>42.284</td>
</tr>
<tr>
<td>CHECK</td>
<td>.950</td>
<td>.217</td>
<td></td>
<td>.588</td>
</tr>
</tbody>
</table>
### TABLE 3
REDUCED FORM ESTIMATION OF LOAN RATE EQUATION

<table>
<thead>
<tr>
<th>Variable</th>
<th>Low Income Sample</th>
<th>High Income Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>t-stat.</td>
</tr>
<tr>
<td>CONSTANT</td>
<td>11.26</td>
<td>4.28</td>
</tr>
<tr>
<td>SEX</td>
<td>-.86</td>
<td>-1.86</td>
</tr>
<tr>
<td>BLACK</td>
<td>3.09</td>
<td>3.02</td>
</tr>
<tr>
<td>RENT</td>
<td>.83</td>
<td>.66</td>
</tr>
<tr>
<td>COLLAG</td>
<td>2.61</td>
<td>2.53</td>
</tr>
<tr>
<td>CHECK</td>
<td>.03</td>
<td>.04</td>
</tr>
<tr>
<td>INC</td>
<td>-.31E-04</td>
<td>-.24</td>
</tr>
<tr>
<td>DIV</td>
<td>-1.07</td>
<td>-.79</td>
</tr>
<tr>
<td>SOUTH</td>
<td>-1.25</td>
<td>-1.15</td>
</tr>
<tr>
<td>NEAST</td>
<td>.07</td>
<td>.06</td>
</tr>
<tr>
<td>NCENT</td>
<td>-1.33</td>
<td>-1.25</td>
</tr>
<tr>
<td>TRES</td>
<td>-.07</td>
<td>-.34</td>
</tr>
<tr>
<td>URB</td>
<td>.14</td>
<td>.65</td>
</tr>
<tr>
<td>LA</td>
<td>-.17E-03</td>
<td>-1.17</td>
</tr>
<tr>
<td>NP</td>
<td>-.05</td>
<td>-.25</td>
</tr>
<tr>
<td>DATT</td>
<td>.06</td>
<td>.29</td>
</tr>
<tr>
<td>RESCH</td>
<td>.22</td>
<td>.17</td>
</tr>
<tr>
<td>AGE</td>
<td>.10E-02</td>
<td>.03</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Low Income</th>
<th>High Income</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOBS</td>
<td>253</td>
<td>323</td>
</tr>
<tr>
<td>S.E.</td>
<td>5.09</td>
<td>3.79</td>
</tr>
<tr>
<td>F Statistic</td>
<td>1.69 (.049)</td>
<td>1.18 (.278)</td>
</tr>
</tbody>
</table>

(significance levels in parentheses)
price discrimination within the low income group but do not practice discrimination within the high income subsample. The null hypothesis that individual characteristics do not affect the interest rate charged may be rejected at the 5% level for the low income subsample and accepted at all standard significance levels for the high income subsample.

These results give stronger support to the idea that lenders practice price discrimination in the low income subsample. The results also indicate that lenders are very sensitive to changes in σ at low σ levels and relatively insensitive to σ when σ is high. The differences may be explained by demand-related factors or some missing variables, such as the competitive structure of the banking industry across regions. If the competitive structure were different across regions, however, the regional dummies should be significant. None of the regional dummies is significant in either of the two regressions.

Although a supply curve for credit has not been identified, it may be possible to determine whether price discrimination occurs by examining which of the
individual characteristics have the most significant effect on interest rates in the low-wealth subsample. The only variables that have a statistically significant effect on the rates that low-wealth individuals face are BLACK and COLLAG. The coefficients on the race and collection agency variables are 3.09 and 2.62 respectively, implying that individuals with these characteristics pay significant interest rate premiums (when they are lent credit). These are the only coefficients that are significant at the 5% level. Although none of the other coefficients are statistically significant, most of them are of the expected sign.

The results above imply that lenders discriminate on the basis of BLACK and COLLAG since it is unlikely that these variables have a significant effect on one's demand for credit. The collection agency variable may be a signal that one likes to borrow, but a very high percentage of the entire sample borrowed over the four year interview period. It is doubtful that the people who have had bills turned over to collection agencies borrow so much more than other people with similar characteristics that 262 basis point premiums are
warranted.

Kain and Quigley (1972) present evidence that blacks pay higher per unit prices in the housing market than comparable whites. They argue that since blacks face higher housing prices, their demand for substitute goods such as durables may increase. The evidence in Table 3 indicates that they also pay higher prices in the consumer loan market, so it is not clear whether they should have a higher or lower demand for consumer durables than comparable whites. It is hard to believe that demand factors alone explain the over 300 basis point premium that blacks pay in the consumer loan market.

Part of the BLACK premium may be explained by indirect screening on the part of lenders. If lenders do not locate their offices in the black community, then blacks will not have as many sources of credit as whites, meaning that they face a less competitive market structure than whites. Also, because blacks may have fewer sources of convenient credit, they may not have as much information about the interest rates that they face as white people do. This does not explain why they
would overestimate, instead of underestimate, the rates that they would face. The indirect screening hypothesis also implies that blacks have lower, not higher demands for credit than whites. In summary, the indirect screening hypothesis implies that blacks pay a higher cost in applying for loans than whites and thus apply less than whites.

The weight of the evidence would indicate that lenders practice price discrimination on low income individuals, but it must still be emphasized that demand-related factors introduce confusion in the interpretation of the results. More detailed data sets would be needed to disentangle the supply and demand factors with any precision.

The reduced form high-wealth subsample results are substantially different from the low-wealth results. The null hypothesis that all of the individual characteristic coefficients are zero (in the high-wealth subsample) cannot be rejected at any significance level. The evidence supports the claim that the individual characteristics of high-wealth individuals are not priced in the consumer loan market.
All but one of the variables are insignificant and many have the wrong signs in the high-wealth regression. The one variable that has a statistically significant impact on the interest rate is DATT. The attitudes towards debt variable ranges from 1 to 5, with 5 representing the least favorable attitude towards taking on debt. The coefficient on this variable is .325, indicating that individuals with poorer attitudes towards taking on debt expect to pay higher interest rates.

The direction of causation between the interest rate and attitude toward debt is unclear. People may have bad attitudes because they face higher interest rates. Conversely, people may believe that they face higher rates precisely because they have bad attitudes towards taking on debt. Although the relationship between the interest rate and attitudes variables is unclear, we can still conclude that some survey bias has been eliminated by the inclusion of this variable.

The regression results reported above may be relatively close to estimates of a loan supply curve because the survey question ("What interest rate do you face in the personal loan market") elicited responses
about how suppliers of credit behaved. The signs on almost all of the coefficients are consistent with a supply side story: variables that increase default risk have positive coefficients and variables that decrease default risk have negative coefficients.

While the econometric results seem plausible, there are weaknesses in the research design. There may be errors in variables bias and truncation bias. The existence of an interest rate ceiling will cause the coefficients estimated by OLS to be inconsistent. The error term will have a truncated normal distribution rather than a full normal distribution when there is an interest rate ceiling. The truncation bias may not be particularly severe, however; the mean interest rate ceiling on a 1000 dollar loan (the size that people were queried about in the interest rate question) was 27.31 in 1970 for a 27 state sample (Paxson (1985)). The mean interest rate faced by the low-wealth subsample was 11.31, so there may not be enough limit observations to bias the OLS coefficients. Barth et al. (1983) found OLS and Tobit interest rate regressions to yield very similar results.
VIII. Conclusion

The results in this paper suggest that individuals face price discrimination in consumer loan markets. Lenders seemingly price certain individual characteristics that are the strongest signals of default probability.

The results in this paper are tentative because (1) reduced form, not structural form, equations were estimated so structural interpretations cannot be applied to the results and (2) the effect of rationing due to interest rate ceilings was not accounted for in the regressions. The extent of the bias due to interest rate ceilings may not be large, but this is a matter for further investigation.
APPENDIX

RATIONING MODEL

The critical assumption that prevents rationing in the model presented in this paper is that \( \sigma \) is observable to the lender. If \( \sigma \) is unobservable to the lender a rationing equilibrium may be constructed.

Suppose that \( \sigma \) is unobservable and distributed uniformly within the borrowing population:

\[
f(\sigma) = \frac{1}{\Gamma}, \quad 0 \leq \sigma \leq b \quad (1)
\]

The mean expected return from serving a class of observationally indistinguishable borrowers is given by

\[
E(\pi) = \frac{\Gamma}{\int_{\sigma^*}^{\Gamma} \sigma E(\pi_i) \, d\sigma / \Gamma}
\]

where \( \sigma^* \) is the minimum level of \( \sigma \) that is necessary to induce an individual to borrow, equal to \( \rho \min^2 / \delta (b - y_1) \) and \( E(\pi_i) \) is given equation (14) in the text, the expected return to the lender from serving a member of the observationally indistinguishable class \( i \).
Equilibrium in the loan market is defined by

\[
\frac{\partial}{\partial \Gamma} \int_{\sigma^*}^{\Gamma} \sigma E(\pi_1) d\sigma / \Gamma = 0 \tag{3}
\]

Equation (3) defines the maximum of the \( E(\pi) \) function for a particular class of borrower. Solving (3) for the interest rate \( r \) yields the equilibrium interest rate \( r^* \). This is identical to the \( r^* \) depicted in Figure 1. A closed form solution for the interest rate cannot be found from equation 3, but it is still possible to determine the effects of all of the exogenous variables on the interest rate by using the implicit function theorem.

The equilibrium described by (3) implicitly assumes that there are no rate ceilings. A few more assumptions are needed to determine the equilibrium if we assume that there are legal loan rate ceilings. In such situations the actual rate charged is usually equal to the rate ceiling.

Finally, it should be noted that interest rates are
still monotonically increasing in one's credit riskiness under one mild assumption. Suppose that $E(\pi^1) > E(\pi^2)$ at all interest rate levels; then $r^1 < r^2$ always. This result suggests that the qualitative influence of individual characteristics on interest rates is not significantly altered by the presence of rationing. It seems that the presence of rationing will shift some, but typically not all, of a lender's response to credit risk from price discrimination to rationing.
FOOTNOTES

(1) In Appendix 1 I explain how the empirical model developed below may be altered to explain interest rate differentials when there is rationing as well. The model in the appendix is useful only to the extent that rationed and unrationed borrowers may be distinguished—it is not a test of rationing.

(2) The level of income protected from bankruptcy varies across states. Typically, bankrupt individuals are allowed to keep the level of income that could be generated by working full time at the minimum wage plus a fixed percentage of income over this amount. There is relatively wide variation, however, in the amounts that individuals may keep. Barth et al. (1983) find that the degree to which wages may be garnished has a significant effect on the loan rate that an individual faces.

(3) There are numerous costs associated with defaulting that are not captured by the $U(y_{min})$ term. There are psychic costs and losses in one's credit rating that may decrease a defaulter's utility level. These costs are not included in the formal model.

(3a) It is technically possible for an individual to borrow more than $(b - y_{min})/r$, but such a possibility could not be supported in a loan market equilibrium because lenders would not have an incentive to lend such amounts at any interest rate.

(4) This is a very weak assumption since we only demand that the borrower has an infinitesimally positive probability of repayment. An individual can always play a lottery at minimum cost, thus ensuring a positive (albeit extremely small) probability of repayment. Since there is always a chance that an individual may reap a windfall, it is reasonable to assume that the probability of repayment is always positive.

(5) If the borrower successfully recognizes that $\partial r/\partial B$ is not equal to zero, he will lower his desired borrowing level. The comparative statics results would
not be altered by this more accurate assumption about borrower behavior. Paxson (1985) developed a maximization model similar to to the one presented here in which it is assumed that borrowers do recognize that $\frac{\partial r}{\partial B}$ is not equal to zero. She does not, however, employ this result in her final form regression model.

(6) If $d$ becomes extremely large, the utility function approaches linearity. An individual with a linear utility function consumes entirely in the first period or entirely in the second period, so the solution for $B$ in such a case is necessarily an endpoint solution, $-\gamma_1$ or $(b-y_{\text{min}})/r$.

(7) This is an extreme simplification, but a highly useful one. Complex state regulations determine exactly how much an individual must pay when he files for bankruptcy. Since the focus of this study is on the effects of individual characteristics on loan rates (not legal variables on loan rates) I will not formally model the legal variables.

(8) This is an entirely reasonable assumption: most finance institutions (and retail firms that lend heavily to individuals) use sophisticated credit scoring models that employ individual characteristics.

(9) That is, the sign of $\frac{\partial r}{\partial x_i}$, the effect of the $i$th individual characteristic on the interest rate may remain unchanged in the presence of rationing. In Appendix 1 a rationing equilibrium is developed. The signs of the $\frac{\partial r}{\partial x_i}$ terms may be the same as they are in the nonrationing equilibrium under certain assumptions about the loan market. The derivatives are not the same when the interest rate equals the interest rate ceiling set by law. In that case the interest rate is not affected by individual characteristics.
REFERENCES


850-872.


Weber, E. "Installment Credit for Autos: A Literature

CREDIT CONSTRAINTS AND THE TIME PATH OF ASSET HOLDINGS OVER THE LIFECYCLE
I. Introduction

This paper addresses two issues regarding the time path of asset holdings over the life-cycle. First, is the time path of asset holdings over the life-cycle hump-shaped? Under most reasonable assumptions about the time path of earnings, the shape of utility functions and length of life, the time pattern of net wealth holding over the life-cycle should be hump-shaped. Recent empirical research, however, has not corroborated this notion. Second, how do individual-specific variables, especially those related to credit constraints, affect wealth holding at a point in time? In the first essay of this dissertation, variables related to credit constraints were found to have a significant impact on the interest rates that individuals face in loan markets. This paper attempts to determine the degree to which such household-specific variables (especially those related to one's credit status) affect the time profile of asset holdings. Their inclusion may substantially improve the model because (1) in theory, liquidity
constraints should affect one's asset holding level and a significant percentage of the population is fully or partially liquidity constrained and (2) researchers have not been completely successful in explaining the time path of asset holdings using life-cycle models that have not accounted for credit constraints.

Numerous researchers have been unable to confirm the proposition that asset holdings are hump-shaped. Mirer (1979) and Brittain (1978) concluded that wealth was a strictly increasing function of age. Blinder, Gordon and Wise (1980) solved for the level of wealth holding at a point in time by an optimizing consumer and concluded that the data provided almost no support for the life-cycle hypothesis and the corresponding hump shape of asset holdings. Dicks-Mireaux (1982) corrected for cross-sectional differences in permanent income to demonstrate that the wealth-age profile is hump-shaped. King and Dicks-Mireaux (1982) accounted for the wealth patterns of only the top percentage of the entire sample, however, claiming that there is a group of people "...for whom the simple life-cycle model is inappropriate...because of capital market constraints."
Consumer maximization theory suggests that liquidity constrained individuals should have lower net asset holdings than unconstrained individuals. In a model with three goods (nondurable, durable and financial), an individual constrained in all loan markets will typically substitute both financial and nondurable goods for the durable goods. The net effect of the constraint is to lower net wealth because of the partial substitution towards nondurables.

Many studies have modeled the effects of age on asset holdings, but none has analyzed the effects of variables related to credit constraints on asset holdings; the model used in this paper simultaneously determines the effects of credit constraints and age on individual asset holdings. This methodology has two important advantages over previous research designs: (1) Instead of removing individuals who may be constrained from the sample it explicitly accounts for cross-sectional differences in individual credit market status and (2) it minimizes specification error due to the exclusion of household-specific terms.

The major empirical results in this study may be
summarized as follows. First, the time path of wealth holding over the life-cycle seems to be hump-shaped, but not in a statistically significant manner. Almost all of the piecewise linear and quadratic terms in the net wealth regression are statistically insignificant at standard significance levels. This result is at variance with those reported in previous studies. Most studies have found wealth to be either increasing or hump-shaped in age in a statistically significant manner. In this paper, neither of these two hypotheses is completely supported. Second, variables that indicate one's credit market status have the greatest impact on wealth holding. This result corresponds closely to the finding in the first essay that credit market related variables have the greatest impact on loan rates that individuals face in consumer loan markets.

The structure of this paper may be summarized as follows. A theoretical model of the wealth-age relationship is presented in section II, the empirical model is developed and discussed in section III, data construction and selection is reported in section IV, the econometric results are presented in section V and the
conclusion is presented in section VI.

II. Consumer Theory

a. In this section I analyze (1) the determinants of the time path of asset holdings when an individual is unconstrained and then carefully delineate how these relations are altered by liquidity constraints. The fundamental determinants of the wealth-age relationship will be analyzed first within the context of a multiperiod perfect certainty model. The assumption of perfect certainty is somewhat unrealistic in wealth accumulation models since the decision to save depends on one's expected lifetime, but the perfect certainty model does generate many of the same qualitative results as models that assume uncertainty about time of life.

The tastes and budget constraints of a representative consumer will be analyzed first. An individual is assumed to have a flow of income in time t of $E_t$ and an initial endowment of wealth, $A_0$. The individual chooses a consumption path, $C_t$, to maximize lifetime utility, which is given by
\[
\sum_{t=0}^{T} b(t) U(C_t) + B(A_T) \tag{2}
\]

where \(U(C_t)\) is the one-period utility function, \(b(t)\) is the discount factor and \(B(A_T)\) is the utility of terminal assets. \(b(t)\) is specified as

\[
b(t) = 1/(1+\rho)^t \tag{3}
\]

and \(U(C_t)\) and \(B(A_T)\) are assumed to be isoelastic:

\[
U(C_t) = C_t^{1-\delta}/1-\delta, \quad B(A_T) = bA_T^{1-\delta}/1-\delta \tag{4}
\]

In a world of certainty with perfect capital markets the only constraint on the maximization of lifetime utility is the lifetime budget constraint:

\[
\sum_{t=0}^{T} C_t/(1+r)^t + A_T/(1+r)^T = A_0 + \sum_{t=0}^{T} E_t/(1+r)^t = A_0 + Y_0 \tag{5}
\]

where \(r\) is the rate of interest (assumed constant over time) and \(Y_0\) is lifetime discounted earnings.

Adapting the maximization plan to a household consisting of \(N_t\) adults when the head of the
household is age t leads to the family utility function:

$$N_t U(C_t/N_t)$$

(5a)

Maximizing lifetime utility subject to (5) leads to the first order conditions:

$$C_t = (N_t/N_0) C_0 (1+g)^t$$

(6)

$$A_T = \beta (C_0/N_0)$$

where $1+g = (1+r/(1+\rho))^{1/\delta}$ and $\beta = (b)^{1/\delta} (1+r)^{T/\delta}$. The level of terminal assets, $A_T$, is primarily determined by the $\beta$ parameter. A high $\beta$ implies a strong bequest motive. Blinder et al. (1980) assumed that $\beta$ was strictly a function of the number of children in the household, but it is clear that numerous other household-specific variables may affect $\beta$.

The budget constraint from time t forward implies that the sum $A_t + Y_t$ must be equal to the present value of future consumption plus the planned bequest. Thus:
\[ A_t + Y_t = \sum_{s=t}^{T} \frac{C_s}{(1+r)^{s-t}} + \frac{A_T}{(1+r)^{T-t}} \]  

(7)

Combining equations (5), (6) and (7) yields the final form relation for \( A_t + Y_t \):

\[ A_t + Y_t = \frac{\sum_{s=t}^{T} (1+\phi)^s N_s + \beta_s}{\sum_{s=0}^{T} (1+\phi)^s N_s + \beta_s} * (A_0 + Y_0)(1+r)^t \]  

(8)

where \( 1+\phi = 1+g/1+r \) and \( \beta^* = \beta(1+r)^{-T} \). The denominator in (8) is the number of adult equivalent years of consumption in the family's life cycle (appropriately discounted, accounting for any trends in desired consumption and allowing for planned bequests through the \( \beta \) term). The number of adult equivalent years of consumption remaining for the household when the head of the household is age \( t \) is given by the numerator in (8). Equation (8) states that the fraction of total lifetime resources still available is equal to the fraction of adult equivalent years of life still remaining.
The shape of the time path of asset holdings over the life cycle is primarily determined by \( \phi \) and \( \beta^* \). The steepness of the time profile of asset holdings is positively related to \( \beta \): the higher the taste for bequests, the higher should one's asset holdings be at a point in time. \(^4\) The \( \beta^* \) term should also be related to a number of household-specific factors. For example, \( \beta^* \) is positively related to the number of children in the family--planned bequests should increase with the number of children in the household.

The steepness of the time profile of asset holdings is positively related to \( \phi \). This result makes intuitive sense, given the functional form of \( \phi \):

\[
\phi = (1+r)/(1+\rho)^{1/\delta}/1+r - 1
\]

\( \partial \phi/\partial \delta < 0 \), \( \partial \phi/\partial \rho < 0 \) and \( \partial \phi/\partial r \) is uncertain

The signs of the derivatives above all make intuitive sense. For example, a high discount rate, \( \rho \), implies a low regard for future consumption and a corresponding low level of wealth holding throughout the life-cycle (and, as the model is constructed above, a relatively flat time
profile of asset holdings).

For the most part, household-specific variables affect the time profile of asset holdings in an ambiguous manner. For example, a household head’s education level may positively or negatively affect $\rho$, $\delta$ and $\beta^*$: no clear theoretical relation has been established between educational levels and these variables in the past. Similar caveats apply to many other household-specific variables (except for the number of children in the household, which has been extensively treated at the theoretical level).

b. The solution for asset holdings derived in (8) was made under the assumption that the household was not liquidity constrained. It is not possible to derive a general closed form solution for asset holdings when credit constraints are present because there are an infinite number of ways that an individual may be credit constrained over his lifetime—he may be fully or partially constrained for any number of years. A few general propositions may, however, be made about the asset holdings of a constrained household. The most important proposition is that a household constrained in
period \( \tau \) will, ceteris paribus, have lower asset holdings in period \( \tau \) than if it had been unconstrained. This result holds in a three good world (durables, nondurables and a financial good) because the household, unable to purchase the utility maximizing level of durables, will partially substitute towards nondurables and thus reduce asset holding levels. This result holds true for most, but not all utility functions.

III. Test Design

a. Ideally, we would like to estimate a version of equation (8) to determine the shape of asset holdings over the life cycle, but there are two obstacles to such a procedure. First, Blinder, Gordon and Wise (1980) estimated just such an equation but generated poor results. They concluded that (1) the model explained very little of the cross-sectional variation in savings behavior, (2) the critical parameters of the life cycle model were very poorly identified and in some cases implied nonsensical relations (for example, negative bequests were implied by the model) and (3) the data was consistent with the life cycle model only if it was
assumed that people's utility functions shift systematically by age in such a way as to produce low consumption levels late in life. Second, equation (8) only applies to households that are unconstrained in personal loan markets—it is therefore unsuitable for describing the asset holding behavior of a sample that contains constrained households.

The model of asset holding adopted in this paper is derived from the general reduced form relation specified by:

\[ \ln(\frac{W_i}{Y_i}) = f(A_i) + \ln(X_i) + e_i \]  

\(W_i/Y_i\): ratio of assets or net worth to permanent income

\(A_i\): The age of the head of the household

\(X_i\): a vector of observable household-specific variables which influence the wealth-age relationship (through both tastes and credit constraints).

\(e_i\): a white noise error term representing unobservable variables and deviations of household preferences from the mean.

The reduced form approach is adopted in this paper because (1) it affords precision in measuring the
wealth-age relationship and (2) it provides flexibility in extending the analysis to account for household-specific factors and (3) little success has been achieved in structural form estimations. There are alternative strategies that may better explain household preferences and constraints, but such strategies typically lead to imprecise parameter estimates.

The hypotheses to be tested are: (1) \( \partial f/\partial A_i \) is positive for values of \( A \) up to the age of retirement and negative thereafter and (2) \( \partial l/\partial x_i \) is equal to zero for each element in the \( X_i \) vector. Several steps must be taken, however, to make equation (10) estimable. First, the functional forms for both \( f \) and \( l \) must be defined. Second, a workable model of permanent income determination must be developed, since permanent income enters the left hand side of (10). These issues will be covered in turn.

The \( f \) function must be flexible enough to test the hypothesis that the pattern of wealth holding over the life-cycle is hump-shaped. The \( f \) function is assumed to be a piecewise linear function of age, implying a constant rate of accumulation within each age bracket.
The linear formulation is restrictive, but most asset holding time profiles (for both constrained and unconstrained households) are better captured by the piecewise linear formulation than by a quadratic formulation, for example.\(^6\) The \(l\) function is assumed to be linear in all of the elements in the \(X_i\) vector. The following dummy variables are used to implement the piecewise linear formulation for \(f\):

\[
\begin{align*}
    d_{1i} &= 1 \text{ if } A_i < 30, \text{ zero otherwise} \\
    d_{2i} &= 1 \text{ if } 30 < A_i < 40, \text{ zero otherwise} \\
    d_{3i} &= 1 \text{ if } 40 < A_i < 50, \text{ zero otherwise} \\
    d_{4i} &= 1 \text{ if } 50 < A_i < 60, \text{ zero otherwise} \\
    d_{5i} &= 1 \text{ if } 60 < A_i < 75, \text{ zero otherwise}
\end{align*}
\]

The corresponding pieces to be used in the final form regression are:

\[
\begin{align*}
    V_{1i} &= d_{1i}(A_i - 15) + 15 \sum_{j=2}^{6} d_{1i} \\
    V_{2i} &= d_{2i}(A_i - 30) + 10 \sum_{j=3}^{6} d_{2i}
\end{align*}
\]
\[ V_{3i} = d_{3i} (A_i - 40) + 10 \sum_{j=4}^{6} d_{3j} \]

\[ V_{4i} = d_{4i} (A_i - 50) + 10 \sum_{j=5}^{6} d_{4j} \]

\[ V_{5i} = d_{5i} (A_i - 60) + 15d_{5i} \]

\[ V_{6i} = d_{5i} (A_i - 60)^2 + 225d_{6i} \]

The second step that must be taken in order to implement equation (10) is to develop a measure of permanent income for each household in the sample. Permanent income is assumed to differ from current income by an age-specific term:

\[ \ln(E_{it}) = \ln(Y_{it}) + g(A_{it} - A^*) + e_{it} \quad (11) \]

where

\[ \ln(E_{it}): \text{the log of earnings for individual } i \text{ in year } t \]

\[ \ln(Y_{it}): \text{the log of permanent income for individual } i \]

\[ A^*: \text{a reference age by which permanent income is defined} \]

\[ e_{it}: \text{a random white noise error} \]
A model of permanent income determination is needed to make equation (11) operational. The basic equation describing permanent income for each household is:

\[ \ln(Y_{it}) = X_{it} b + t_i + h(A_i) \]  

(12)

\( X_{it} \): a vector of observable variables for individual \( i \) that affect permanent income, such as education

\( t_i \): an unobservable individual effect

\( h(A_i) \): a cohort effect designed to correct for intergenerational differences in technical skills and training.

Combining equations (11) and (12) yields the final form expression for current earnings:

\[ \ln(E_{it}) = X_{it} b + g(A_{it} - A^*) - h(A_i) + t_i + e_{it} \]  

(13)

The error term, \( t_i + e_{it} \), has zero mean and a variance of \( \sigma_t^2 + \sigma_e^2 \). The \( t_i \) and \( e_{it} \) terms are assumed to be uncorrelated with one another.

A number of the terms in (13) must be given specific functional forms or be assigned values in order to make

71
(13) an estimable equation. Both the g and h functions need specific functional forms and $t_i$ must be assigned a value since a panel data set would be needed to separately identify the $t_i$ and $e_{it}$ terms.

The g function is approximated by a cubic function of age. The h function is constructed following a methodology identical to that employed by King and Dicks-Mireaux (1982). The h function is assumed to be described by a piecewise linear function so that at the age of 45 it has a value of zero. A value for $t_i$ may be generated from the information contained in the residuals from the estimated earnings equation, $t_i + e_{it}$. The minimum variance estimator of $t_i$ (using the information in $t_i + e_{it}$) is

$$t_i^* = \alpha(t_i + e_{it}) \quad (14)$$

where

$$\alpha = \frac{\sigma^2_t}{\sigma^2_t + \sigma^2_e} \quad (14a)$$

Permanent income may be calculated by using the information from estimating the following equation:

$$\ln Y_i = X_i b^* - h(A_{it}) + t_i^* \quad (15)$$
where \( t^*_i \) is the product of \( \alpha \) and the residual from the earnings equation and the estimated value of \( b, b^* \), is used. The estimate of permanent income in equation (15) is consistent because the error in this estimate is uncorrelated with the information used to construct the estimate (and thus the estimate itself). The error in the estimate is

\[
    u_{it} = \ln Y^*_t - \ln Y_i = \alpha e_{it} - (1-e_{it})t_i \quad (16)
\]

\[
    E(u_{it} * \ln Y^*_i) = E(u_{it}(t_i + e_{it})) \quad (17)
\]

If \( \alpha \) is constructed as in equation (14a) above then (17) will be zero.

The value of \( \alpha \) is imposed in the estimated model because \( \sigma^2_r \) and \( \sigma^2_e \) cannot be estimated without a complete panel data set. The value of \( \alpha \) is determined using non-sample information. Researchers using similar data sets have generated a relatively narrow range of estimates for \( \sigma^2_r \) and \( \sigma^2_e \). Lillard and Willis (1979) used the Panel Study of Income Dynamics data set to examine 1,144 male heads of household for seven years. They
estimated an earnings equation with both small and large sets of explanatory variables and generated values for \( \alpha \) of .471 and .606. Lillard (1977) and Lillard and Weiss (1979) obtained similar results. The value of \( \alpha \) is assumed to be .5 in constructing estimates of permanent income.

The final form equation to be estimated is:

\[
\ln(W_i/Y_i) = a_0 + \sum_{j=1}^{6} a_j V_{j1} + b'X_i + u_i
\]  

The coefficients from \( a_1 \) to \( a_4 \) measure the growth rate of the wealth to permanent income ratio in the four youngest age ranges. The \( a_5 \) and \( a_6 \) coefficients are critical in determining the growth rate of the wealth to permanent income ratio after retirement. If the time profile of the wealth to permanent income ratio is hump-shaped then \( a_6 \) must be negative. The \( X_i \) vector contains liquidity constraint variables and other individual specific variables which may affect the growth path of the wealth to permanent income ratio.

The household-specific variables are constrained to
affect the intercept but not the age growth coefficients in equation (18).

Credit constraints typically shift entire sections of the time path of asset holdings so the piecewise linear structure of equation (18) is ideal for capturing such changes.

IV. Data Selection and Construction

The wealth measure used in this study includes the value of cash, deposits, bonds, stocks and shares, savings plans, other financial and non-financial assets, motor vehicles and household durables less certain household debts. The net wealth measure excludes the value of social security and pension plans.

The variables in the $X_i$ vector are race (RACE), time in current residence (TRES), number of people in the household (NP), years of work experience (WKEX), educational level (ED), and two dummies reflecting whether the household head has ever been refused credit (REFCRED) or had a bill sent to a collection agency (COLL). Table 1 lists the variables used and their definitions.
TABLE 1
DATA DESCRIPTION AND DEFINITION

COLLAG\_i\: A dummy variable that takes the value of one if the individual has ever had a bill turned over to a collection agency, zero otherwise.

REFCRED\_i\: A dummy variable that takes on the value of one if the individual has ever been refused credit, zero otherwise.

RACE\_i\: A dummy variable taking on the value of one if the individual is black, zero otherwise.

TRES\_i\: Time spent at current residence. The variable ranges from 1 to 9, with 9 representing the shortest stay at the current residence. Increases in the TRES\_i variable imply decreases in the time spent at the current residence.

W\_i\: Net assets at the beginning of the period. The sum of amounts held in checking and savings accounts, bonds, stocks, other financial assets and durables.

EARN\_i\: The combined earned incomes of the household head and wife.

NP\_i\: Number of people in the household.

AGE\_i\: Age of household head.

WKEX\_i\: The years of work experience (in current occupational field) of the household head.

ED\_i\: A variable representing how well educated the household head is. The variable ranges from 1 to 9, with 9 representing a graduate degree.
The NP variable should be positively related to the number of children or people in the household. A positive bequest motive should be reflected by higher wealth holdings at a point in time. The WKE and ED variables have been found to be positively correlated with earnings, but are not necessarily related to wealth levels.

The REFCRED and COLL variables should give a partial picture of the individual's status in the credit markets. In the first essay, I found that individuals with each of these characteristics paid significantly higher interest rates in consumer loan markets. These variables may therefore be useful in explaining movements in the wealth to permanent income ratio since that ratio may be affected by credit constraints.

This study utilizes the "Consumer Durables and Installment Debt" ICPSR data set. This data set contains information on income wealth and personal characteristics for 1436 households. All observations with complete data available were used, reducing the total sample to 1281 observations.
V. Econometric Results

Table 2 shows the distribution of net assets by age. The most important result in this table is the low levels of wealth held by a large number of the households in the sample. Over 20 percent of the entire sample has net wealth lower than 2000 dollars. This is in line with the results reported in Diamond and Hausman (1980), who found that approximately 20 percent of the men aged 45 to 59 reported wealth less than 1000 dollars in 1966 (in a sample of over 5,000 males). King and Dicks-Mireaux (1982) found that almost 30 percent of their sample of over 10,000 Canadian households reported net wealth of less than 3000 dollars in 1976.

The low net wealth figures reported here do not contradict the life cycle hypothesis because net wealth, as defined in this study, excludes the value of both private and public pension rights. These results are consistent with the Blinder et al. (1980) result that households systematically underprovide for future generations.

The net wealth and net wealth to permanent income
### TABLE 2

**DISTRIBUTION OF WEALTH**

<table>
<thead>
<tr>
<th>DECILE</th>
<th>NET WEALTH</th>
<th>GROSS WEALTH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MEAN</td>
<td>STD. DEV.</td>
</tr>
<tr>
<td>1</td>
<td>110</td>
<td>837</td>
</tr>
<tr>
<td>2</td>
<td>1595</td>
<td>309</td>
</tr>
<tr>
<td>3</td>
<td>2632</td>
<td>266</td>
</tr>
<tr>
<td>4</td>
<td>3582</td>
<td>292</td>
</tr>
<tr>
<td>5</td>
<td>4545</td>
<td>285</td>
</tr>
<tr>
<td>6</td>
<td>5697</td>
<td>384</td>
</tr>
<tr>
<td>7</td>
<td>7317</td>
<td>558</td>
</tr>
<tr>
<td>8</td>
<td>10402</td>
<td>1222</td>
</tr>
<tr>
<td>9</td>
<td>17562</td>
<td>3061</td>
</tr>
<tr>
<td>10</td>
<td>50660</td>
<td>27998</td>
</tr>
</tbody>
</table>

79
ratios by age group are reported in Table 3. The two
statistics present slightly different pictures of wealth
holding by age. The ratio suggests that wealth holding
reaches a maximum at some point between the ages of 60
and 64 and then declines dramatically in the 65 to 69 age
range, implying that the time path of asset holdings is
hump-shaped. The net wealth statistics indicate that
wealth holding reaches a maximum at a much earlier age,
perhaps in the 45 to 49 age range. The net wealth to
permanent income ratio is superior to net wealth as a
measure of financial well-being because households with
identical asset holdings but different permanent income
levels should not be considered identical.

The results from estimating regression (18) are
reported in Table 4 (Estimates of (18) after outliers are
removed are presented in Table 5.) The inverse Mills
ratio (derived from a probit regression) is included to
correct for the sample selection bias that occurs because
the model is only estimable for positive net wealth
levels. The effects of the age group coefficients on
the wealth to permanent income ratio will first be
discussed and then the effects of the household-specific
<table>
<thead>
<tr>
<th>AGE</th>
<th>NET WEALTH/PERM INC</th>
<th>NET WEALTH</th>
</tr>
</thead>
<tbody>
<tr>
<td>15-24</td>
<td>.568</td>
<td>6224</td>
</tr>
<tr>
<td>25-29</td>
<td>.838</td>
<td>11118</td>
</tr>
<tr>
<td>30-34</td>
<td>1.008</td>
<td>11781</td>
</tr>
<tr>
<td>35-39</td>
<td>.921</td>
<td>8703</td>
</tr>
<tr>
<td>40-44</td>
<td>.968</td>
<td>10224</td>
</tr>
<tr>
<td>45-49</td>
<td>1.258</td>
<td>11713</td>
</tr>
<tr>
<td>50-54</td>
<td>1.296</td>
<td>10805</td>
</tr>
<tr>
<td>55-59</td>
<td>1.211</td>
<td>10380</td>
</tr>
<tr>
<td>60-64</td>
<td>1.427</td>
<td>7053</td>
</tr>
<tr>
<td>65-69</td>
<td>.556</td>
<td>5889</td>
</tr>
</tbody>
</table>
### Table 4

**Dep. Variable:** Wealth to Permanent Income Ratio

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Estimated Coefficient</th>
<th>t-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CONSTANT</strong></td>
<td>-0.25033</td>
<td>-0.56888</td>
</tr>
<tr>
<td><strong>V1</strong></td>
<td>1.88087e-002</td>
<td>0.61564</td>
</tr>
<tr>
<td><strong>V2</strong></td>
<td>1.43515e-002</td>
<td>0.98597</td>
</tr>
<tr>
<td><strong>V3</strong></td>
<td>4.09271e-002</td>
<td>2.93336</td>
</tr>
<tr>
<td><strong>V4</strong></td>
<td>-7.20070e-004</td>
<td>-4.43985e-002</td>
</tr>
<tr>
<td><strong>V5</strong></td>
<td>6.33363e-002</td>
<td>0.60340</td>
</tr>
<tr>
<td><strong>V6</strong></td>
<td>-4.95327e-003</td>
<td>-0.61797</td>
</tr>
<tr>
<td><strong>RACE</strong></td>
<td>-0.29901</td>
<td>-2.59066</td>
</tr>
<tr>
<td><strong>COLL</strong></td>
<td>-7.51587e-002</td>
<td>-0.67102</td>
</tr>
<tr>
<td><strong>NP</strong></td>
<td>-5.99308e-002</td>
<td>-2.27098</td>
</tr>
<tr>
<td><strong>REFCREDS</strong></td>
<td>-0.19895</td>
<td>-1.87789</td>
</tr>
<tr>
<td><strong>ED</strong></td>
<td>-1.05523e-002</td>
<td>-0.32520</td>
</tr>
<tr>
<td><strong>TRES</strong></td>
<td>-2.65656e-002</td>
<td>-1.36343</td>
</tr>
<tr>
<td><strong>WKEX</strong></td>
<td>-6.41868e-003</td>
<td>-1.17476</td>
</tr>
<tr>
<td><strong>INV. MILLS</strong></td>
<td>-6.35546</td>
<td>-2.93350</td>
</tr>
</tbody>
</table>

**Number of Observations** 1058  
**R-squared** 8.00044e-002  
**Corrected R-squared** 6.76555e-002  
**Sum of Squared Residuals** 1.11649e+003  
**Standard Error of the Regression** 1.03463  
**Mean of Dependent Variable** -0.59623

82
### Table 5

**Outliers Removed**

**Dep. Variable:** Wealth to Permanent Income Ratio

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Estimated Coefficient</th>
<th>t-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONSTANT</td>
<td>-9.17336e-002</td>
<td>-0.21860</td>
</tr>
<tr>
<td>V1</td>
<td>5.15433e-003</td>
<td>0.17689</td>
</tr>
<tr>
<td>V2</td>
<td>2.10889e-002</td>
<td>1.51421</td>
</tr>
<tr>
<td>V3</td>
<td>3.89743e-002</td>
<td>2.91075</td>
</tr>
<tr>
<td>V4</td>
<td>-3.16002e-003</td>
<td>-0.20421</td>
</tr>
<tr>
<td>V5</td>
<td>0.13947</td>
<td>1.36313</td>
</tr>
<tr>
<td>V6</td>
<td>-1.60020e-002</td>
<td>-1.83577</td>
</tr>
<tr>
<td>RACE</td>
<td>-0.27075</td>
<td>-2.46139</td>
</tr>
<tr>
<td>COLL</td>
<td>-5.26993e-002</td>
<td>-0.49327</td>
</tr>
<tr>
<td>NP</td>
<td>-5.71014e-002</td>
<td>-2.24944</td>
</tr>
<tr>
<td>REFCRED</td>
<td>-0.25403</td>
<td>-2.49002</td>
</tr>
<tr>
<td>ED</td>
<td>-1.56090e-002</td>
<td>-0.49775</td>
</tr>
<tr>
<td>TRES</td>
<td>-3.04102e-002</td>
<td>-1.63005</td>
</tr>
<tr>
<td>WKEX</td>
<td>-4.10572e-003</td>
<td>-0.78461</td>
</tr>
<tr>
<td>INV. MILLS</td>
<td>-7.12911</td>
<td>-3.37871</td>
</tr>
</tbody>
</table>

Number of Observations: 1035
R-squared: 0.10098
Corrected R-squared: 8.86416e-002
Sum of Squared Residuals: 9.79433e+002
Standard Error of the Regression: 0.97991
Mean of Dependent Variable: -0.64564
variables on the wealth to permanent income ratio will be analyzed.

The coefficient estimates on the $V_j$ variables indicate that net wealth holding over the life-cycle (corrected for differences in permanent income) increases over the life-cycle and decreases after retirement. The negative $V_6$ coefficient implies a hump-shaped growth path with a maximum reached at age 66.4 ($60 - a_0 / 2a_0$). Despite the plausible parameter estimates, the only growth rate variable that is significantly different from zero is $V_3$. The t-statistics thus imply that the life-cycle hypothesis is not supported by the data. The null hypothesis that all of the $V_j$ coefficients are zero can be rejected at the 5 percent level but not at the 1 percent level (using an F test).

An anomaly in the results is the negative sign on the $V_4$ coefficient. A possible explanation for this result is that the people aged 50 to 60 may have been adversely affected by beginning their working lives at the start of World War II.

The annual rate of decumulation of wealth is given by the derivative of (18) with respect to age. For ages 60
to 75 the rate of wealth decumulation is given by \(-[a_0 + 2a_0(A_{lt} - 60)]\). The approximately 2% decumulation rate at the age of 75 implied by the model is consistent with simulations conducted by Davies (1980); Davies assumed constant relative risk aversion and uncertainty about date of death and found an annual decumulation rate of 2%. The decumulation rates are higher in this model than in King, Dicks-Mireaux (1982), who generated rates between 1/2 and 1 percent. The results do not change dramatically when the outliers are removed from the sample: the implied maximum of the time path of wealth holdings occurs at the age of 64.4.

The econometric results reported above contradict the findings of Mirer (1979) and Brittain (1978) who found that wealth holdings monotonically increase over the lifecycle. Dicks-Mireaux (1982) reported similar results using a similar methodology, but their estimates of the growth rate coefficients were significant more often than those reported in this paper. Conversely, the results in Tables 4 and 5 are somewhat different from those developed by Blinder et al. (1980), who concluded that the data gave almost no support to the life cycle
hypothesis. The difference in method may explain the relative success of the results in this paper and in King, Dicks-Mireaux (1982) compared to the Blinder et al. (1980) results.

While the age growth variables explain some of the variation in asset holdings, most of the variation in asset holdings is explained by the household-specific variables. This result suggests, but does not prove, that credit constraints play a critical role in determining the shape of asset holdings over the life cycle. This result may also explain why numerous researchers have been unable to successfully explain the time path of asset holdings using narrowly defined life-cycle models.

The variables that are closely associated with one's credit market status have the greatest impact on one's wealth holdings. In the first essay the REFCREd, COLL and RACE variables were found to have strong and statistically significant effects on the interest rates that individuals face. The present study shows that these variables have a corresponding negative impact on the amount of wealth held by individuals. In particular, the
REFCRED variable has a strong and statistically significant negative effect on wealth holding, suggesting that both past and current credit constraints may influence wealth holdings. The REFCRED coefficient estimate implies that people who have been refused credit have net wealth to permanent income ratios that are approximately 20 percent lower than the mean ratio in the entire sample. This result is consistent with the notion that individuals constrained in durable loan markets will partially substitute nondurable for durable goods, thus reducing their holdings of illiquid assets.

The size of the RACE dummy coefficient implies that blacks have substantially lower holdings of net wealth than the rest of the sample. This difference in wealth holding may be explained by a wide variety of social and economic forces, but part of such an explanation includes the constraints that blacks face in loan markets, as documented in the first essay of this dissertation and Avery (1983).

Table 6 reports summary statistics for the full sample and for the subsample of individuals who have been refused credit or had a bill sent to a collection agency.
### TABLE 6

**SUMMARY STATISTICS**

<table>
<thead>
<tr>
<th></th>
<th>&quot;CONSTRAINED SAMPLE&quot;</th>
<th></th>
<th>&quot;UNCONSTRAINED SAMPLE&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MEAN</td>
<td>STD. DEVIATION</td>
<td>MEAN</td>
</tr>
<tr>
<td>RACE</td>
<td>.1255</td>
<td>.3320</td>
<td>.0758</td>
</tr>
<tr>
<td>NP</td>
<td>3.91</td>
<td>1.91</td>
<td>3.703</td>
</tr>
<tr>
<td>EARN</td>
<td>11058</td>
<td>14655</td>
<td>11719</td>
</tr>
<tr>
<td>ED</td>
<td>4.34</td>
<td>1.94</td>
<td>4.39</td>
</tr>
<tr>
<td>AGE</td>
<td>44.69</td>
<td>11.06</td>
<td>43.22</td>
</tr>
</tbody>
</table>
The "constrained" subsample was, on average, older, less educated and had a lower income than the "unconstrained" subsample. Unsurprisingly, the "constrained" subsample contained a significantly larger percentage of blacks than the "unconstrained" subsample. In summary, the regression estimates suggest that credit market constraints (or at least the variables that are associated with such constraints) lead to lower asset holdings. These conclusions must be considered tentative because they are based on interpretations of reduced form regressions.

VI. Conclusion

The time path of asset holdings was examined using a cross-sectional data set of over 1000 households. Regression estimations suggest that, contrary to recent evidence, the pattern of wealth holding over the life cycle is consistent with the life-cycle hypothesis: net wealth holding was hump-shaped over the life cycle, as the life cycle hypothesis suggests that it should be. Although the coefficient estimates suggest a plausible time path of asset holdings, few of the coefficient
estimates in the net wealth regressions were statistically significant, suggesting that there are a wide variety of motives for saving within the population.

Individual-specific variables were found to explain more of the cross-sectional variation in the net wealth to permanent income ratio than the age variables. Credit constraint variables in particular had significant effects on the wealth to permanent income ratio. This result corresponds closely to the results in the first essay of the dissertation. In that chapter, race and a few credit constraint variables were the only individual-specific variables to affect the interest rate that individuals faced in loan markets. In this essay such variables were found to have the greatest influence on wealth to permanent income ratios among a selection of individual-specific variables. It would be natural to infer that the influence of these variables on one's credit market standing was directly transmitted to one's net wealth position. However, it is difficult to make such an inference since both the credit market and net wealth regressions carry reduced form, not structural, interpretations. If structural interpretations could be
derived, it is possible but unlikely that the results of this study would be reversed.
(1) The failure of Blinder et al. (1980) to generate meaningful results may not be due to weaknesses in the life cycle model per se. The model of the lifecycle that they estimated is highly nonlinear and overparameterized by most standards. It's possible that the estimation success of models in this field is inversely related to the complexity of the equations estimated.

(2) The model presented in this section was initially developed by Blinder et al. (1980).

(3) A case could be made for the notion that educational levels are negatively related to one's discount factor. One indirect and imprecise way of determining this is to estimate the correlation between educational levels and one's participation in activities that are hazardous to one's health.

(4) It should be emphasized that this result holds for most but not all utility functions.

(5) Blinder et al. (1980) may have failed to generate meaningful bequest predictions because they were unable to get precise estimates of parameters that purported to measure (1) the wealth elasticity of bequests and (2) the effect of the number of children on asset accumulation.

(6) The piecewise linear formulation may best characterize the time path of asset holdings because most individuals go through discrete stages in their working lives that are best captured by the discrete, time-specific piecewise linear variables. King and Dicks-Mireaux (1982) were the first researchers to employ the piecewise linear methodology presented in this section.

(7) In theory, credit constraints should affect the intercept and the age growth coefficients. That is, credit constraints should affect both the level and growth rate (within an age group) of asset holdings at a
point in time.

(8) A ten percent yearly depreciation rate on the household durable stock was assumed on purchases made up to three years prior to the survey year. No data was available on durables purchases made over three years prior to the survey year. It was therefore assumed that this unobserved durable stock constituted half of the total durable stock. This is a conservative assumption since the average level of durable purchases in the three years prior to the survey year and the survey year are relatively low. Also, all of the financial variables are expressed in nominal terms. An inflation correction should not substantially affect the results because the left-hand side of the final regression, $\frac{W_i}{Y_i}$, remains unchanged by such a correction.

(9) King and Dicks-Mireaux (1982) estimated regressions with and without pension expected pension value variables. They found expected pension wealth to have a statistically significant effect on asset holdings. They also found that the inclusion of expected pension wealth did not significantly alter the other coefficient estimates, suggesting that their exclusion from the model leads to relatively minor specification error.

(10) A similar correction was made in the regressions used to estimate permanent income.
REFERENCES


706-733.


CORPORATE LOAN DEFAULT AND THE PRICING OF RISK IN THE CORPORATE BOND MARKET
I. Introduction

This paper addresses the issue of whether the bond market efficiently utilizes all available information about company default prospects in setting bond prices. If the bond market accurately analyzes all information about a firm, then accounting and stock market information should not improve the discrimination between defaulting and nondefaulting firms (i.e., the conditional probability of default, given the firm's bond yield, should not be affected by any accounting or stock market information). The concept of market efficiency discussed here is somewhat different than the standard definition of efficiency in financial markets. Market efficiency in this context simply means that no information apart from bond yield spreads should increase the accuracy of a corporate loan default prediction.

The issues addressed in this paper have not, by and large, been addressed in the bankruptcy prediction literature. Most of the empirical literature has focused on the development of optimal prediction models and the corresponding determination of which accounting ratios
best predict bankruptcy (notable exceptions are Ohlson (1980) and Zmijewski (1984)). Unfortunately, there has been little agreement on the correct methods to use and even less agreement on which accounting ratios best predict bankruptcy. There are two reasons why there has been such disagreement in the literature. First, bankruptcy regressions are almost never based on a well-specified theory of bankruptcy (primarily because few such theories exist), leading to ad hoc variable selection techniques and models that have very poor out-of-sample prediction rates. Wilcox (1976) constructed empirical models based on the gambler’s ruin model but these models performed worse than atheoretical bankruptcy prediction models in discriminating between bankrupt and nonbankrupt firms. Second, bond market-based measures of default/bankruptcy risk are never employed. Some researchers such as Altman et al. (1977) have employed stock market volatility variables, but such measures give only a vague indication of bankruptcy risk.

Previous studies have attempted to determine the value of accounting information in predicting bankruptcy because accounting information presumably summarizes the
financial condition of a firm. Market participants, however, are informed of many factors beyond the published accounting information of a firm. Variables that explain the sources of supply for a firm, the trend in its market share, its financial flexibility and access to capital markets are almost never included in bankruptcy prediction models. Bond yields should presumably reflect such information about a firm. The only circumstances under which accounting information should have predictive value in a prediction model that includes bond yield spreads is if bond market participants misread the signals contained in the financial statements of firms.\(^{(2)}\)

This study circumvents the difficulties involved with accounting-based prediction models by grounding the discrimination of defaulting from nondefaulting firms in the theory of market efficiency. In this essay I develop a market-based model of bond default risk. The effectiveness of the market-based measure of risk in predicting corporate loan defaults is then compared to numerous accounting and stock market measures using a multivariate logit model. The test design employed in
this paper is superior to those used in previous studies because (1) it is the first to determine the marginal predictive content of accounting and stock market information within the context of a model based on a formal theory and (2) it is the first to use a bond market-based default indicator in loan default regressions.

The major results of this essay may be summarized as follows: (1) the bond market does not efficiently incorporate all accounting and stock market information into bond prices, (2) the size of dividend distributions (as a percentage of total assets) has significant marginal predictive content in all logit default regressions and (3) the rejection of market efficiency is robust to adjustments in the bond yield spread, sample time period, covenant valuation and number of years prior to default that bond market and accounting information is used. In summary, it seems that bond market participants do not use an appropriate default prediction model.

The structure of the essay is as follows: a probabilistic model of loan default is developed in section II, timing issues are discussed in section III, a
bond yield spread model is developed in section IV, data selection and definition is covered in sections V and VI, the test design is presented in section VII, econometric results are presented in section VIII and the conclusion is in section IX.

II. Default Probability Model

The primary hypothesis examined in this paper is that the conditional probability of loan default, given a firm's bond yield, is unaffected by any other information. Stated more precisely:

\[ P_{it,t+j}(D_{i,t+j}|B_{it}) = P_{it,t+j}(D_{i,t+j}|B_{it},A_{it}) \quad \forall i,\forall j \]

(1)

where

\( D_{i,t+j} \) = a dummy variable that takes on the value of one if firm i defaults on a loan in period \( t+j \), zero otherwise.

\( B_{it} \) = a vector of variables that summarize the bond market's assessment of firm i's default risk in time period t.
\( A_{it} \) = a vector consisting of accounting and stock market variables about firm \( i \) in period \( t \).

\( P_{t,t+j} \) = The probability as of time period \( t \) that firm \( i \) will default on its loans in time period \( t+j \).

Equation (1) states that accounting or stock market information has no marginal effect on the probability of default at a point in time, given the bond market's assessment of default risk.

The hypothesis that accounting information has no marginal effect on the conditional default probability, given the bond market's default probability assessment, may be tested with an equation of the form:

\[
D_{i,t+j} = f(B_{it}, A_{it})
\]  

(1a)

where \( j \) is typically chosen to equal one year. If the bond market is efficient, \( \partial f / \partial A_{it} = 0 \).

A strength of this test is that the choice of the bond yield spread as a right hand side variable is grounded in a formal theory of bond market efficiency. Conversely, the choice of variables used to test the
alternative hypothesis (that market efficiency does not hold) does not have to be based on a formal theory. The power of the test does increase, however, if the ratios that best predict loan default are used.

III. Timing Issues

There are two major issues that must be addressed in the implementation of the efficiency test: (1) the determination of optimal lag lengths between (a) the default date and the posting of the bond prices and (b) the posting of bond prices and the release of other information such as 10-K reports and (2) the determination of the bond market’s assessment of a firm’s default probability. These issues will be discussed in turn.

The determination of the optimal lag between the release of bond market (and accounting) data and loan default dates has been the subject of numerous discussion (Ohlson, (1980)). The length of this lag affects the power of bond yield spreads (and potentially accounting information) in discriminating between defaulting and nondefaulting firms. If the lag is made short enough
there should be few errors in discriminating between the two types of firms. The length of this lag should not affect the test for efficiency, however; no accounting or stock market information should have an effect on the conditional default probability at any point in time, given the yield spread.

The second important lag is that between the posting of the bond prices and the release of accounting data. The length of this lag should affect the power of accounting ratios in discriminating between defaulting and nondefaulting firms. If the lag is reduced to approximately zero (i.e., bond price listing coincides with release date of accounting information) the accounting ratios and stock market data may help discriminate between defaulters and nondefaulters because the bond market may not immediately incorporate such information into the price of bonds.

The lag between the release date of accounting information and the posting of bond prices used in this study is approximately a month (end of March versus end of April). The relatively long lag is chosen because the goal of this study is not to analyze the ability of the
bond market to quickly incorporate all information into bond prices. The goal is to determine if the bond market efficiently utilizes accounting and stock market information in assessing long-run default probabilities. The lag used in this study (between the release date of accounting information and the posting of bond prices) is therefore relatively long, leading to a bias towards accepting the hypothesis of bond market efficiency. The bias is potentially significant in those cases where the firm releases its first quarter earnings report in the interim period between its release of its 10-K report and the bond price listing.

IV. Bond Yield Spread Model and Selection of Bond Market Data

In order to develop an estimable version of equation (1a), the elements in both the $B_{it}$ and $A_{it}$ vectors must be defined. These elements will be discussed in turn. There are a number of bond market based measures of default risk. A commonly used measure of default risk is the yield spread between the corporate bond and a Treasury bond of corresponding coupon and maturity. The
yield spread between the riskless and risky bond should provide a relatively accurate measure of the bond market's assessment of default risk. There are factors apart from default risk, however, that may affect the corporate yield (and therefore the yield spread). A simple model of the yield spread that accounts for such characteristics unrelated to default is given by

\[ \text{YSPRED}_i = V(C_i) + f(M_i) + g(\text{cou}_i, \text{mat}_i) + \varepsilon_i \quad (2) \]

where

\( \text{YSPRED}_i \) = The yield spread between corporate bond \( i \) in time period \( t \) and a matching Treasury bond of similar coupon and maturity.

\( \varepsilon_i \) = The default premium or measure of default risk incorporated in corporate bond \( i \) in time \( t \).

\( V(C_i) \) = The value of the call option on corporate bond \( i \) in time \( t \).

\( f(M_i) \) = A function of the marketability or liquidity of bond \( i \) in time \( t \).

\( g( ) \) = A function of the difference in coupon and maturity between corporate bond \( i \) and the corresponding Treasury bond.

The portion of the yield spread that is not explained
by the right-hand side variables in equation (2) represents the default premium. The list of right-hand side variables in (2) may not be complete, however, because corporate bonds often have a variety of covenants that are not explicitly accounted for in equation (2). This issue is discussed further in the econometric results section.

The yield spread is constructed by matching the corporate bonds in the sample with Treasury bonds of similar coupon and maturity. The g function captures any differences in yield that are due to maturity or coupon differences; the g function is given by

\[ g(\text{cou}_i, \text{mat}_i) = \alpha_1(\text{cou}_i - \text{cou}_{tb}) + \alpha_2(\text{mat}_i - \text{mat}_{tb}) \]  

The signs of \( \alpha_1 \) and \( \alpha_2 \) will partly depend on whether the Treasury yield curve is upward or downward sloping.

An analytical solution to the value of a call option with a variable striking price on a coupon bond is not available. The call value is a complex function of (1) the time to maturity of the bond, (2) sequence of call prices and (3) variance of the corporate bond price. I
will model the value of the call option with a general quadratic form:

$$C(V_i) = \alpha_3 R_{tb} + \alpha_4 R_{tb}^2 + \alpha_5 \text{mat}_i + \alpha_6 \text{mat}_i^2 + \alpha_7 \text{mat}_i * R_{tb} \quad (4)$$

The quadratic approximation may be relatively accurate in this instance because a call value is most nonlinear when it is nearly in the money and none of the call options on the bonds in this sample are nearly in the money. Finally, since the time to maturity and sequence of call prices are relatively stable across firms, most of the variation in call values may be explained by the overall movement of interest rates in the economy, the $R_{tb}$ variable.

A marketability measure may explain a fraction of the yield spread since investors are willing to pay a premium for liquidity. Bonds that are relatively illiquid should have relatively higher yields in order to compensate holders for the difficulty in selling such bonds. The dollar value of the bonds outstanding is used as a proxy for the marketability of the bonds:
\[ f(M_1) = \alpha_8 \text{OUTS}_i \]  

A precise measure of marketability would be difficult to construct for corporate bonds, so the simple proxy of the amount outstanding is the preferred measure.

The final form equation for the yield spread is given by

\[ YS_i = \alpha_0 + \alpha_1 (\text{cou}_i - \text{cou}_{lb}) + \alpha_2 (\text{mat}_i - \text{mat}_{lb}) + \alpha_3 R_{lb} + \alpha_4 R_{lb}^2 + \alpha_5 \text{mat}_i + \alpha_6 \text{mat}_i^2 + \alpha_7 \text{mat}_i R_{lb} + \alpha_8 \text{OUTS}_i + \epsilon_i \]

\[ = X_i'\alpha + \epsilon_i \]  

Although equation (6) represents a fairly crude model of corporate bond yield spreads, it should serve the purposes of this study well. Casual empiricism suggests that the variables in equation (6) should not account for more than 10 to 20 percent of the variation in yield spreads.

V. Selection and Definition of Accounting Variables

Since theory suggests that no variable apart from the
information contained in the bond yield spread should have a statistically significant effect in equation 1, any variable should be considered an acceptable member of the $A_{it}$ vector. Correspondingly, there are no compelling reasons why some accounting ratios should be chosen over others, mainly because nobody has developed a convincing theory of corporate bankruptcy. Some researchers have employed stepwise regression methods to select those variables which seem to have the most explanatory power. The stepwise procedure may yield good in-sample classification rates but the tests of statistical significance are biased and the out of sample classification rates are typically lower than the in-sample rates (Zmijewski (1984)).

The financial ratios used in this study have (1) been most successful in previous corporate bankruptcy studies and (2) cover the major operating financial characteristics of firms. This method of data selection is not ideal, but the test of bond market efficiency does not hinge on the choice of variables in the $A_{it}$ vector.

The accounting ratios and stock market data used are:
ROA = Net income divided by total assets.

STOCK = The standard deviation of the firm's monthly stock return over the two years prior to default.\(^5\)

CUMPROF = Retained earnings divided by total assets.

CURRENT = Current liabilities divided by current assets.

LEV = Total debt as a percentage of total assets.

SIZE = The log of the firm's total tangible assets.

DIV = Total dividend distributions on common and preferred stock divided by total assets.

INTCOV = Operating income divided by total interest payments.

The most successful of these ratios in bankruptcy studies have been the DIV, STOCK and CUMPROF variables. Using stepwise methods, univariate F-tests and conditional deletion tests, Altman (1983) found that the CUMPROF variable was always the most important variable in explaining bankruptcy. Altman used a paired sample design so his results must be interpreted with care.

The LEV variable plays a critical role in
options-based formulations of the yield spread. The LEV variable has not, however, performed particularly well in empirical bankruptcy prediction models (see Altman (1983)). Altman, Haldeman and Narayan (1977) found a measure similar to STOCK improved bankruptcy prediction. The yield spread on a corporate bond should still be superior to STOCK as a default predictor, however, (even though they are both market-based) for the simple reason that the stock volatility should be reflected in the yield spread.

VI. Data

The corporate bond price and yield data used for this study are retrieved from issues of the Standard and Poor's Bond Guide. The S&P prices are compiled from a variety of sources. All of the prices are quotes from either the New York and American bond exchanges or brokers.

The Standard and Poor's prices (whether from the exchange or brokers) represent either the sale or bid price on the last trading day of the month (end of April for firms with fiscal year ends). The prices used in this study should therefore reflect all available
accounting information on firms at the time of the price quote (regardless of whether they are sale or bid prices).

Most corporate bond trading is done directly through brokers, not through the exchanges, so it is possible, but unlikely (because of arbitrage opportunities) that the exchange prices are a poor reflection of value in the corporate bond market. Conversely, because there is not a single price on most corporate bonds, there are potential biases in virtually any approach to corporate bond pricing. This issue is discussed further in the econometric results section.

The accounting data for the nondefaulting firms are obtained from the Compustat research file. There are 189 nondefaulting and 29 defaulting firms in the sample. The sample of non-defaulting firms was constructed by randomly choosing firms from Standard and Poor's Bond Guides for the years 1980 to 1985. The non-defaulting firms were then assigned to a particular year between 1980 and 1985 in order to have a similar percentage of firms in each year as the defaulting sample. Each firm had to satisfy the following criteria in to be included
in the sample. First, a firm had to have bond prices available for two consecutive years because logit regressions using one and two year prior to default data are estimated. Second, a firm had to be in a non-financial industry group (as determined by the Standard and Poor's classification system). Financial services firms were excluded from the sample because they have a number of different structural characteristics than other firms. Third, a firm had to have complete data available on the relevant bond market and accounting figures employed in this study. If the sample of unchosen firms without complete data differs from the sample of firms that do have complete data then the statistical tests will be inconsistent. Zmijewski (1984) corrected his logit results for the inconsistency caused by the complete data criterion, but the correction had a negligible effect on the econometric results.

The sample of defaulting firms was taken from Altman and Nammacher (1985) and Standard and Poor's Bond Guides over the 1980-1985 sample period. A firm was considered in default if its bonds were listed as D (in default) by Standard and Poor's or if it filed for protection from
its creditors under Chapter XI. Data for the defaulting firms was taken from the Compustat file and firm 10-K reports.

VII. Test Design

The equation used to test the market efficiency hypothesis is a logit model of the form

\[ D_{it+1} = \alpha'A_{it} + \delta'YSPRED_{it} - \beta'X_{it} + u_{it+1} \]  \hspace{1cm} (7)

where

\[ D_{it+1} = \begin{cases} 
1 & \text{if the firm has defaulted on its corporate debt obligations in period } t+1. \\
0 & \text{if the firm has not defaulted on its corporate debt obligations as of period } t+1. 
\end{cases} \]

and

\( A_{it} \): a 8 x 1 vector of the accounting and stock market variables defined in section V.

\( YSPRED_{it} \): the yield spread between firm i's bond and the corresponding Treasury bond in period t, which represents the default premium in this equation, given the presence of the \( X_{it} \) variables.

\( \alpha \): an 8 x 1 vector of coefficients
\( \beta \): an 8 x 1 vector of coefficients.

\( u_{it} \): a random white noise error term.

Under certain restrictive assumptions, the hypothesis that accounting and stock market information has no marginal effect on the probability of default is simply a test of \( \alpha_1 = \alpha_2 = \ldots = \alpha_8 = 0 \).\(^9\) The most important assumption underlying this test is that the accounting and bond market information affect the conditional probability in a linear manner.

The nonlinear cross-equation constraints implied by equation 6 are not employed in the equation 7 hypothesis test because it is difficult to correctly implement the estimation and testing procedure when the constraints are imposed (with little potential gain in efficiency). Constrained versions of the test are discussed in the econometric results section, but accurate standard errors are not calculated.

The focus of this study is the efficiency test, not the development of a premier prediction model, so no observations will be withheld for a hold-out sample. The power of the efficiency test is increased because there
are more observations but the classification ability of the logit model can no longer be accurately tested with out of sample data.

VIII. Econometric Results

Table 1 reports the means and standard deviations of ratios and bond yield spreads of defaulting and nondefaulting firms both one and two years prior to default. Virtually all of the statistics have the expected magnitudes: the defaulting firms have weaker balance sheets, higher stock return volatilities and higher yield spreads than the nondefaulting firms. The differences in the means are almost all statistically significant at the 1 percent level (for both one and two years prior to default). Also, all of the defaulting firms' measures of financial strength deteriorate between two and one years prior to default. The yield spread for defaulting firms is over three times as great as the spread for nondefaulting firms; similar comparisons apply to a number of the other financial ratios, including the DIV, INTCOV and CUMPROF variables.

The multivariate logit estimation results using
# Table 1

**Profile Analysis**

**One Year Prior to Default**

<table>
<thead>
<tr>
<th></th>
<th>Defaulting Firms</th>
<th></th>
<th>Non-Defaulting Firms</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
<td><strong>Std. Dev.</strong></td>
<td><strong>Mean</strong></td>
<td><strong>Std. Dev.</strong></td>
<td></td>
</tr>
<tr>
<td><strong>CURRENT</strong></td>
<td>.6857</td>
<td>.3774</td>
<td>.5820</td>
<td>.2209</td>
</tr>
<tr>
<td><strong>LEV</strong></td>
<td>.4551</td>
<td>.1598</td>
<td>.2207</td>
<td>.1121</td>
</tr>
<tr>
<td><strong>ROA</strong></td>
<td>-.0528</td>
<td>.1286</td>
<td>.0445</td>
<td>.0658</td>
</tr>
<tr>
<td><strong>DIV</strong></td>
<td>.0051</td>
<td>.0067</td>
<td>.0223</td>
<td>.0133</td>
</tr>
<tr>
<td><strong>CUMPROF</strong></td>
<td>.0713</td>
<td>.1446</td>
<td>.3214</td>
<td>.1551</td>
</tr>
<tr>
<td><strong>SIZE</strong></td>
<td>6.0886</td>
<td>1.2653</td>
<td>7.5495</td>
<td>1.2388</td>
</tr>
<tr>
<td><strong>STOCK</strong></td>
<td>.1290</td>
<td>.0311</td>
<td>.0923</td>
<td>.0307</td>
</tr>
<tr>
<td><strong>YS/PRED</strong></td>
<td>5.1227</td>
<td>2.6046</td>
<td>1.6529</td>
<td>1.2481</td>
</tr>
</tbody>
</table>

**Two Years Prior to Default**

<table>
<thead>
<tr>
<th></th>
<th>Defaulting Firms</th>
<th></th>
<th>Non-Defaulting Firms</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
<td><strong>Std. Dev.</strong></td>
<td><strong>Mean</strong></td>
<td><strong>Std. Dev.</strong></td>
<td></td>
</tr>
<tr>
<td><strong>CURRENT</strong></td>
<td>.6633</td>
<td>.2408</td>
<td>.5648</td>
<td>.2029</td>
</tr>
<tr>
<td><strong>INT/COV</strong></td>
<td>2.4806</td>
<td>1.5174</td>
<td>8.5528</td>
<td>7.5811</td>
</tr>
<tr>
<td><strong>LEV</strong></td>
<td>.3986</td>
<td>.1643</td>
<td>.2196</td>
<td>.1096</td>
</tr>
<tr>
<td><strong>ROA</strong></td>
<td>.0004</td>
<td>.0690</td>
<td>.0510</td>
<td>.0717</td>
</tr>
<tr>
<td><strong>DIV</strong></td>
<td>.0061</td>
<td>.0070</td>
<td>.0223</td>
<td>.0131</td>
</tr>
<tr>
<td><strong>CUMPROF</strong></td>
<td>.1345</td>
<td>.1110</td>
<td>.3298</td>
<td>.1480</td>
</tr>
<tr>
<td><strong>SIZE</strong></td>
<td>6.0368</td>
<td>1.2094</td>
<td>7.4850</td>
<td>1.2548</td>
</tr>
<tr>
<td><strong>STOCK</strong></td>
<td>.1265</td>
<td>.0325</td>
<td>.0936</td>
<td>.0365</td>
</tr>
<tr>
<td><strong>YS/PRED</strong></td>
<td>4.2434</td>
<td>2.1004</td>
<td>1.7126</td>
<td>1.3087</td>
</tr>
</tbody>
</table>
accounting and bond market information from one and two years prior to default are presented in Tables 2 and 3 respectively. The logit estimations for both one and two years prior to default indicate that the bond market does not efficiently incorporate all information about a firm's default prospects into bond prices. The null hypothesis that all of the accounting and stock market variable coefficients are zero may be rejected at the 1 percent level for both one and two years prior to default (using the appropriate likelihood ratio test statistic in each case). The efficiency test is probably biased towards accepting the null hypothesis of efficiency because the most recently published quarterly accounting data (prior to the bond price date) are not used for the firms in the sample. It is therefore unlikely that efficiency is rejected because the bond market is slow in reacting to changes in a firm's default risk. Finally, virtually identical results were achieved by estimating the constrained version of equation (7).

The dividend variable (DIV) is most responsible for the rejection of the null hypothesis. It is statistically significant at the 5% level in the one
TABLE 2
LOGIT DEFAULT ESTIMATION
ONE YEAR PRIOR TO DEFAULT

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Estimated Coefficient</th>
<th>t-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>current</td>
<td>1.58197</td>
<td>0.78504</td>
</tr>
<tr>
<td>intcov</td>
<td>-0.24366</td>
<td>-0.85618</td>
</tr>
<tr>
<td>lev</td>
<td>7.29021</td>
<td>2.01238</td>
</tr>
<tr>
<td>roa</td>
<td>-2.03198</td>
<td>-0.46029</td>
</tr>
<tr>
<td>div</td>
<td>-1.86506e+002</td>
<td>-2.21873</td>
</tr>
<tr>
<td>cumprof</td>
<td>-0.60960</td>
<td>-0.15232</td>
</tr>
<tr>
<td>yspred</td>
<td>0.88239</td>
<td>2.53720</td>
</tr>
<tr>
<td>stock</td>
<td>-9.68019</td>
<td>-0.88877</td>
</tr>
<tr>
<td>Δmat</td>
<td>0.51742</td>
<td>2.18782</td>
</tr>
<tr>
<td>Δcou</td>
<td>0.18542</td>
<td>0.57673</td>
</tr>
<tr>
<td>outs</td>
<td>-3.36294e-002</td>
<td>-1.94489</td>
</tr>
<tr>
<td>τb</td>
<td>14.22401</td>
<td>1.44497</td>
</tr>
<tr>
<td>mat</td>
<td>-0.64005</td>
<td>-0.47457</td>
</tr>
<tr>
<td>size</td>
<td>1.11051</td>
<td>1.92131</td>
</tr>
<tr>
<td>mat^2</td>
<td>-7.46923e-003</td>
<td>-0.44034</td>
</tr>
<tr>
<td>τb^2</td>
<td>-0.63172</td>
<td>-1.55085</td>
</tr>
<tr>
<td>τb*mat</td>
<td>7.08991e-002</td>
<td>0.69349</td>
</tr>
<tr>
<td>one</td>
<td>-90.44174</td>
<td>-1.45321</td>
</tr>
</tbody>
</table>

auxiliary statistics at convergence
log likelihood -29.275
log likelihood ratio index .806
number of observations 218
percent correctly predicted 94.95413
chi square test statistic 27.415
(test of α_1=...=α_8= 0)

* significant at the 1% level

120
<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Estimated Coefficient</th>
<th>t-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>current</td>
<td>2.53702</td>
<td>1.51615</td>
</tr>
<tr>
<td>intcov</td>
<td>-0.27379</td>
<td>-1.50746</td>
</tr>
<tr>
<td>lev</td>
<td>4.98423</td>
<td>1.38828</td>
</tr>
<tr>
<td>roa</td>
<td>-3.45176</td>
<td>-0.69840</td>
</tr>
<tr>
<td>div</td>
<td>-1.87478e+02</td>
<td>-2.93117</td>
</tr>
<tr>
<td>cumprof</td>
<td>1.16865</td>
<td>0.35601</td>
</tr>
<tr>
<td>yspread</td>
<td>0.65515</td>
<td>2.58271</td>
</tr>
<tr>
<td>stock</td>
<td>-6.26750</td>
<td>-0.59479</td>
</tr>
<tr>
<td>Δmat</td>
<td>0.41036</td>
<td>1.96280</td>
</tr>
<tr>
<td>Δcou</td>
<td>0.26271</td>
<td>1.09178</td>
</tr>
<tr>
<td>outs</td>
<td>-1.79811e-02</td>
<td>-1.91346</td>
</tr>
<tr>
<td>tb</td>
<td>2.51907</td>
<td>0.55287</td>
</tr>
<tr>
<td>mat</td>
<td>0.12201</td>
<td>0.17445</td>
</tr>
<tr>
<td>size</td>
<td>0.54978</td>
<td>1.28854</td>
</tr>
<tr>
<td>mat²</td>
<td>-2.04162e-02</td>
<td>-1.17590</td>
</tr>
<tr>
<td>tb²</td>
<td>-0.15523</td>
<td>-0.79510</td>
</tr>
<tr>
<td>mat*tb</td>
<td>5.29846e-02</td>
<td>0.78634</td>
</tr>
<tr>
<td>one</td>
<td>-20.23133</td>
<td>-0.74288</td>
</tr>
</tbody>
</table>

**auxiliary statistics**

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>log likelihood</td>
<td>-34.858</td>
</tr>
<tr>
<td>log likelihood ratio index</td>
<td>.769</td>
</tr>
<tr>
<td>number of observations</td>
<td>218</td>
</tr>
<tr>
<td>percent correctly predicted</td>
<td>92.201</td>
</tr>
<tr>
<td>chi square test statistic</td>
<td>36.156</td>
</tr>
<tr>
<td>(test of $a_1=...=a_8 = 0$)</td>
<td></td>
</tr>
</tbody>
</table>
* significant at the 1% level
year regression and at the 1% level in the two year regression. This result is similar to the Gentry et al. (1985) finding that a dividend ratio was the only statistically significant variable (at the 5% level) in a bankruptcy prediction model. The DIV variable should be a more powerful discriminating device than accounting ratios because it reflects management expectations regarding future cash flows—it does not merely reflect current operating conditions. It seems that management sends a strong signal about the financial health of the firm when setting dividend levels. The dividend variable may be particularly useful in discriminating between defaulting and nondefaulting firms because management can misstate accounting numbers but not the size of dividend distributions. Still, the yield spread should supersede the dividend variable in predicting loan defaults because it presumably accounts for the value of the dividend variable.

There are a number of potential explanations for the rejection of the market efficiency hypothesis. A list of explanations might include: (1) The bond yields of the defaulting firms do not reflect the most recent
accounting information because the yields are not formed on the basis of recent trades, (2) the corporate bond yields are mismeasured, (3) neither the raw nor the corrected yield spreads reflect the valuation of bond covenants and (4) the final few defaults (chronologically) among the entire defaulting subsample may have significantly different characteristics than previous defaults and thus skew the statistical results despite the fact that bond market participants use a default prediction model that accurately employs historical default information. These explanations will be considered in turn.

An experiment was conducted to determine the degree to which the results may have been affected by out of date bond yields. The logit default regression (7) was estimated using one year prior to default bond yields and two year prior to default accounting data. This test may help determine whether the non-trading of defaulting bonds (leading to out of date published yields of defaulting bonds) is responsible for the failure of market efficiency. Even if the bond yields are out of date (because of thin trading) they should still reflect
more information than the accounting information released approximately 13 months prior to the bond yield release date.

The logit results using one year ahead bond yields and two year ahead accounting data are reported in Table 4. Surprisingly, market efficiency is still rejected at all standard significance levels (using the log likelihood ratio to test whether all of the $\alpha$ coefficients are equal to zero). The dividend variable DIV is significant at the 5% level and the CURRENT variable is significant at the 10% level. These results imply that non-trading of the defaulting bonds (leading to out of date published yields on defaulting bonds) cannot explain the success of DIV and other variables in improving loan default prediction.

Another potential cause of the rejection of market efficiency is mismeasured bond yields. Because most trading of corporate bonds is over the counter, not through the exchanges, one could argue that broker quotes are the only accurate prices in the market. Potential arbitrage possibilities should limit the differences between exchange and broker prices, but differences may
TABLE 4
LOGIT DEFAULT ESTIMATION
ONE YEAR AHEAD BOND DATA, TWO YEAR AHEAD ACCOUNTING DATA

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Estimated Coefficient</th>
<th>t-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>current</td>
<td>3.70850</td>
<td>1.86381</td>
</tr>
<tr>
<td>intcov</td>
<td>-0.37097</td>
<td>-1.62178</td>
</tr>
<tr>
<td>lev</td>
<td>2.93008</td>
<td>0.86942</td>
</tr>
<tr>
<td>roa</td>
<td>3.65931</td>
<td>0.47915</td>
</tr>
<tr>
<td>div</td>
<td>-1.64212e+002</td>
<td>-2.14671</td>
</tr>
<tr>
<td>cumprof</td>
<td>1.78094</td>
<td>0.48044</td>
</tr>
<tr>
<td>yspread</td>
<td>0.93809</td>
<td>3.15204</td>
</tr>
<tr>
<td>stock</td>
<td>2.79520</td>
<td>0.28326</td>
</tr>
<tr>
<td>Δmat</td>
<td>0.46617</td>
<td>2.35160</td>
</tr>
<tr>
<td>Δcou</td>
<td>0.11286</td>
<td>0.38394</td>
</tr>
<tr>
<td>outs</td>
<td>-1.54784e-002</td>
<td>-1.17076</td>
</tr>
<tr>
<td>tb</td>
<td>12.92801</td>
<td>1.38334</td>
</tr>
<tr>
<td>mat</td>
<td>-1.35378</td>
<td>-1.14696</td>
</tr>
<tr>
<td>size^2</td>
<td>0.56084</td>
<td>1.18919</td>
</tr>
<tr>
<td>mat^2</td>
<td>-5.47694e-003</td>
<td>-0.35264</td>
</tr>
<tr>
<td>tb</td>
<td>-0.62307</td>
<td>-1.57494</td>
</tr>
<tr>
<td>mat*tb</td>
<td>0.12724</td>
<td>1.38987</td>
</tr>
<tr>
<td>one</td>
<td>-74.43911</td>
<td>-1.31417</td>
</tr>
</tbody>
</table>

auxiliary statistics

at convergence
log likelihood
-31.212
log likelihood ratio index
0.793
number of observations
218
percent correctly predicted
92.201
chi square test statistic
23.542*(test of \( \alpha_1 = \ldots = \alpha_6 = 0 \))

* significant at the 1% level
persist because of liquidity differences between the two markets.

To test whether the use of exchange prices had a significant effect on the results, I estimated the logit model using only the defaulting firms with broker quotes and then using only the defaulting firms with exchange quotes. Market efficiency was decisively rejected (at the 1% level) in both the broker and exchange price regressions (and the DIV variable had nearly identical t-statistics in both regressions, 2.38 and 2.47). The rejection of market efficiency may still be attributed to bond yield mismeasurement, just not to the use of exchange prices.

The third explanation for the rejection of market efficiency is that the bond yields do not accurately reflect the value of bond covenants. Casual empiricism suggests that default risk and the number of covenants placed on a bond are positively correlated. Typically, firms with higher default risks are allowed to issue corporate bonds only if covenants restricting the actions of the firm are placed on the bond. If this is the case, the raw yield spread for risky firms will underestimate
their default risk whereas the raw yield spread for firms with a high credit standing will overestimate their default risk. The mean raw yield spread between risky and safe firms is thus probably narrowed by the absence of an accurate covenant valuation. Conversely, even if covenants were accurately valued the efficiency hypothesis may be rejected because the value of the covenants may represent a small percentage of the raw yield spread.

The presence of covenants may also change the behavior of a firm, perhaps altering the efficiency test. Many covenants restrict a firm's financial ratios (such as the ones employed in this essay) to remain above or below a specific limit. These covenants thus increase a firm's incentive to misstate its financial ratios (in order to avoid the penalties associated with the violation of covenants). If the bond market accounts for this adverse incentive, the marginal predictive power of accounting ratios should decrease. The overall effect of this adverse incentive effect therefore is to (1) lower the value of covenants (and thus reduce the bias caused by their non-valuation in the previous tests) and (2)
reduce the power of accounting and stock market information relative to bond yields in discriminating between defaulting and non-defaulting firms.

A crude correction for the valuation of covenants is employed to test whether covenants can partially explain the rejection of market efficiency. I assume somewhat arbitrarily that covenants reduce the yield spread on defaulting firms' bonds by 100 basis points and 0 basis points on nondefaulting firms' bonds. The model is reestimated using a 100 basis point increase in the spreads of the defaulters bonds. The results using the spread adjustment are reported in Table 5 for the one year prior to default regressions (the two year ahead results are similar to the one year results). The use of the adjusted spread does not significantly alter the results: the null hypothesis of market efficiency is still rejected at the 1 percent level and the DIV variable is still significant at the 5 percent level.

Perhaps a more sophisticated valuation of bond covenants could reverse the results reported thus far, but this is doubtful. The covenants probably cannot explain more than ten to twenty percent of the variation
TABLE 5
LOGIT ESTIMATION
CORRECTED FOR COVENANTS

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Estimated Coefficient</th>
<th>t-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>current</td>
<td>2.04089</td>
<td>0.74363</td>
</tr>
<tr>
<td>intcov</td>
<td>-3.76434e-002</td>
<td>-0.10699</td>
</tr>
<tr>
<td>lev</td>
<td>9.98882</td>
<td>1.93944</td>
</tr>
<tr>
<td>roa</td>
<td>-1.64679</td>
<td>-0.28815</td>
</tr>
<tr>
<td>div</td>
<td>-2.78099e+002</td>
<td>-2.36314</td>
</tr>
<tr>
<td>cumprof</td>
<td>-0.64667</td>
<td>-0.11384</td>
</tr>
<tr>
<td>yspred</td>
<td>1.97965</td>
<td>3.11969</td>
</tr>
<tr>
<td>stock</td>
<td>-21.04977</td>
<td>-1.37010</td>
</tr>
<tr>
<td>Δmat</td>
<td>0.77423</td>
<td>2.17171</td>
</tr>
<tr>
<td>Δcou</td>
<td>0.51858</td>
<td>1.03169</td>
</tr>
<tr>
<td>outs</td>
<td>-4.86640e-002</td>
<td>-1.78157</td>
</tr>
<tr>
<td>tb</td>
<td>16.57955</td>
<td>1.27650</td>
</tr>
<tr>
<td>mat</td>
<td>-0.55418</td>
<td>-0.27619</td>
</tr>
<tr>
<td>size</td>
<td>2.00680</td>
<td>2.13373</td>
</tr>
<tr>
<td>mat&lt;sup&gt;2&lt;/sup&gt;</td>
<td>-2.24927e-002</td>
<td>-0.86228</td>
</tr>
<tr>
<td>tb&lt;sup&gt;2&lt;/sup&gt;</td>
<td>-0.74963</td>
<td>-1.38331</td>
</tr>
<tr>
<td>tb*mat</td>
<td>9.40583e-002</td>
<td>0.62228</td>
</tr>
<tr>
<td>one</td>
<td>-1.13088e+002</td>
<td>-1.34428</td>
</tr>
</tbody>
</table>

auxiliary statistics at convergence
- log likelihood: -19.343
- log likelihood ratio index: .872
- number of observations: 218
- percent correctly predicted: 97.247%
- chi square test statistic (test of $a_1=...=a_8=0$): 19.336

*significant at the 1% level
in yield spreads and the results thus far are robust to minor changes in the yield spread.

The fourth explanation for the failure of market efficiency is that dividends may have dramatically affected whether a firm defaulted or not only over the last year or two of the sample period. If this is the case then the bond market should not have been expected to set bond prices according to the size of dividend distributions. For example, if the last four or five defaults (chronologically) are responsible for the rejection of the null hypothesis then the bond market may not be inefficient. It’s more likely, however, that the bond market was surprised by the character of the last four or five defaults and either didn’t believe that the trend would persist or was gradually adjusting to the trend.

To test whether the DIV variable had a late impact on default status, the logit default model was reestimated over the first five years (1980-1984) of the six year sample. The results are reported in Table 6. The results in Table 6 imply that the DIV variable is the most important non-spread variable over the first five years
### TABLE 6

**LOGIT ESTIMATION**

**ESTIMATION PERIOD: FIRST FIVE YEARS OF SAMPLE**

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Estimated Coefficient</th>
<th>t-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>current</td>
<td>0.21032</td>
<td>8.30099e-002</td>
</tr>
<tr>
<td>intcov</td>
<td>-0.20146</td>
<td>-0.61007</td>
</tr>
<tr>
<td>lev</td>
<td>6.17582</td>
<td>1.30680</td>
</tr>
<tr>
<td>roa</td>
<td>-1.01431</td>
<td>-0.19777</td>
</tr>
<tr>
<td>div</td>
<td>-1.51085e+002</td>
<td>-1.70804</td>
</tr>
<tr>
<td>cumprof</td>
<td>-5.61771</td>
<td>-0.99364</td>
</tr>
<tr>
<td>yspread</td>
<td>0.57494</td>
<td>1.64130</td>
</tr>
<tr>
<td>stock</td>
<td>-5.35826</td>
<td>-0.41368</td>
</tr>
<tr>
<td>Δmat</td>
<td>0.53272</td>
<td>2.10730</td>
</tr>
<tr>
<td>Δcoun</td>
<td>0.32519</td>
<td>0.85190</td>
</tr>
<tr>
<td>out5</td>
<td>-2.61859e-002</td>
<td>-1.17764</td>
</tr>
<tr>
<td>tb</td>
<td>15.79372</td>
<td>1.28560</td>
</tr>
<tr>
<td>mat</td>
<td>-0.59415</td>
<td>-0.34092</td>
</tr>
<tr>
<td>size</td>
<td>1.10211</td>
<td>1.42791</td>
</tr>
<tr>
<td>mat²</td>
<td>-9.23146e-003</td>
<td>-0.47080</td>
</tr>
<tr>
<td>tb²</td>
<td>-0.69380</td>
<td>-1.33073</td>
</tr>
<tr>
<td>tb*mat</td>
<td>6.88830e-002</td>
<td>0.51919</td>
</tr>
<tr>
<td>one</td>
<td>-98.18355</td>
<td>-1.27008</td>
</tr>
</tbody>
</table>

auxiliary statistics at convergence initial
log likelihood -22.167 -112.289
log likelihood ratio index .802
number of observations 162
percent correctly predicted 95.061*
chi square test statistic 22.189
(test of $\alpha_1 = \ldots = \alpha_q = 0$)

* Significant at the 1% level

131
of the sample; the DIV variable is significant at the 10 percent level. So, the dividend variable has a statistically significant impact on the discrimination of defaulting from nondefaulting firms prior to the last year of the sample. The rejection of market efficiency therefore cannot be explained by the slow adjustment of the bond market to the importance of dividend distributions.

IX. Conclusion

A model of corporate loan default was developed to determine if the bond market efficiently utilizes accounting and stock market information in assessing corporate debt default probabilities. The null hypothesis that no accounting or stock market information improved the discrimination between defaulting and nondefaulting firms was rejected at the 1 percent level in all reported regressions. The tests were biased against rejecting the null hypothesis so the results may be robust to further perturbations in the test model.

The bond market seems particularly negligent in accounting for differences in dividend distributions.
between firms. It seems that dividend distributions contain a very strong signal about the future prospects of the firm. The dividend variable may be more successful than the accounting ratios because it is easier for management to lie about accounting figures than about dividend sizes.

The model presented in this paper provides a relatively simple description of bond market efficiency and its limitations are numerous. First, the model of corporate bond yield spread determination is crude and does not account for the value of covenants in a sophisticated manner. Second, the corporate bond yields employed may not be accurate barometers of value in the overall corporate bond market. The quotations used may be from brokers that do not do a substantial amount of trading in the bonds or they may be from the bond exchanges, where little corporate bond trading is done. Third, the sample percentage of defaulting firms, 29/218, does not correspond to the population percentage of defaulting firms (less than 1%), potentially leading to biased parameter estimates and standard errors.
FOOTNOTES

1. An abridged list of the studies and their contribution to the literature is: Beaver (1966), introduction of financial ratios to predict bankruptcy, Altman (1968), development of multivariate bankruptcy prediction model, Blum (1974), inclusion of stock market data into prediction models, Ohlson (1980), development of random sampling logit analysis which is superior to the previous multiple discriminant models and Casey and Bartczak (1985), a careful focus on the usefulness of cash flow as an indicator of financial distress.

2. The accounting measures may also have predictive value when the bond market information is mismeasured.

3. Sometimes different Treasury bonds were used to form the one year and two year lagged yield spreads. Ideally, the Treasury and corporate bonds should have identical coupons and maturities. Two bonds of different maturities may have different reactions to information which affects the bond market (even if the two bonds have identical coupons). Also, the Treasury yield curve may differ from year to year, making the measurement of the yield spread a more problematic task.

4. Fisher (1959) found that the elasticity of the yield spread with respect to the dollar value of bonds outstanding was -.29 (significant at the 5 percent level), as the theory predicts.

5. A few of the defaulting firms did not have two years of stock price information available. Such firms were kept in the sample if (1) they had at least one and a half years of price information and (2) they had complete price information in the year preceding default. The bias from using firms with slightly less than two years of price information is most likely minimal.

7. The nondefaulting firms were randomly selected from the "Primary, Supplemental & Tertiary Industrial File" and the "Full Coverage and Over-the-Counter Files".

8. Financial service firms could be included but their operating characteristics are dramatically different from those of industrials. Probably the most important difference from a default theory perspective is that financial services firms can go bankrupt overnight. One can often see the impending signals of trouble ahead for an industrial firm, but a run on a bank can transform a seemingly healthy bank into bankruptcy very quickly.

9. The model is a complete test of the efficiency of the bond market if the set of financial ratios exhausts all potential information about debt default not contained in the bond prices. This is an extremely stringent requirement and is impossible to meet without exhausting the model's degrees of freedom.

10. The two year ahead data is used for these estimations because the logit maximum likelihood routine would not converge for the non-exchange price model. The two year ahead results are useful partly because the one and two year results are similar for almost all of the models.
REFERENCES


BIOGRAPHICAL NOTE


He attended Wesleyan University from 1977 to 1981 and graduated with a Bachelor's degree in Economics in May, 1981. He was awarded the White Prize in Economics and was elected to Phi Beta Kappa in the Spring of 1981.

He was enrolled in the Ph.D. program in Economics at MIT from 1981 to 1987.