1999

Examining agricultural investment

Chad Edward Hart
Iowa State University

Examining agricultural investment

by

Chad Edward Hart

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Co-majors: Agricultural Economics; Statistics

Major Professors: Sergio H. Lence and Alicia L. Carriquiry

Iowa State University

Ames, Iowa

1999

Graduate College
Iowa State University

This is to cerify that the Doctoral dissertation of

Chad Edward Hart

has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.

Co-major Professor

Signature was redacted for privacy.

Co-major Professor

Signature was redacted for privacy.

For the Co-major Program

Signature was redacted for privacy.

For the Co-major Program

Signature was redacted for privacy.

For the Graduate College

DEDICATION

To my wife, Jennifer, and our child,

who is due the same time as this dissertation

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Using two different approaches, the relationship between a firm's investment and its financial variables is examined. Imperfections in the credit market such as asymmetric information have led researchers to explore these relationships. This study incorporates the 5 Cs of lending (character, capacity, collateral, credit rating, and capital) into the farmer's investment decision and explores the impacts of these variables on a basic data set. The data set is composed of 590 Iowa farms that are members of the Iowa Farm Business Association and have reported farm level financial and production data from 1991 to 1995.

The first approach consists of a composite regression model constructed from various elements of traditional investment models and variables representing the 5 Cs. The second approach derives an investment equation from the firm's optimization problem, an Euler equation approach. The 5 Cs of lending are incorporated into the problem through a borrowing constraint.

The composite regression approach is conducted under a Bayesian framework with variable selection and outlier detection components. The results imply strong support for the accelerator model of investment and the inclusion of other relevant variables, among them the value of short-term assets, one of the proxies for the 5 Cs. Another of the proxies, operator age, receives less support. The Bayesian framework with the variable selection and outlier detection components works extremely well.

The Euler equation approach is more problematic. Under the original specification looking at net investment, all models are rejected and the most preferred model is also the most restrictive with symmetric adjustment costs and no financial constraint. Within the financial constraint, only the value of short-term assets and net worth are ever found to be statistically significant. Estimated adjustment costs are either negative or positive but extremely small. The shadow value of external finance is estimated to be around 100 percent. Other formulations, extensions, and reduced form models are explored and similar results are found. Given the mostly negative results from the Euler equation framework, possible reasons for them are reviewed.

CHAPTER 1. INTRODUCTION

There exists a voluminous literature studying the links between finance and investment. Much of this literature is devoted to explore whether there is a link. Early econometric works, such as Tinbergen (1939) and Klein (1951), included financial variables in their investment analyses and found them to be significant. However, Modigliani and Miller (1958, 1961, and 1963) found conditions under which a firm's market value is independent of its capital structure. This implies that the firm is indifferent between financing investment with external or internal funds. After this finding, the development of econometric work linking investment and financial variables came to a virtual standstill. Several recent works have sparked a renewed interest in the subject.

One assumption which leads to the Modigliani and Miller conclusion is the existence of perfect capital markets.[1] However, there are several reasons to believe this is not the case for agriculture. There are transactions costs and tax considerations to take into account. There is an asymmetric information problem between the lender and the farmer as the farmer knows more about his/her probabilities of success and failure than the lender does. Also, numerous studies that included internal financial variables in investment models have found them to be significant. The agricultural credit scoring literature has outlined the attributes most lenders seek in clients. These have been summarized as the "5 Cs": character, capacity (cash flow), collateral, credit rating, and capital (owner equity). If the farmer faces an

[1] The conclusion can also be reached under the less restrictive assumption that firms and investors have the same investment opportunities (Stiglitz, 1969).

additional financial constraint due to imperfect capital markets, the constraint is likely to depend on these 5 Cs.

In this study, we examine the relationship between a firm's investment and its financial variables using two approaches. In the first approach, we form a composite regression model based on several investment models and variables representing the 5 Cs of lending. A Bayesian method is employed to estimate the model within a variable selection, outlier detection framework. In the second approach, we derive an investment equation from the firm's Euler equations.[2] A borrowing constraint is added to the model with the associated multiplier modeled as a function of the firm's financial variables.

The present study adds to the existing literature in several ways. First, most of the literature has focused on the link between investment and a very narrow group of financial variables (in a majority of studies, only one financial variable is included in the model). Our study allows for the possibility of relationships between investment and several financial variables, thus expanding the possible linkages. Second, the Bayesian approach we have chosen focuses on the problem of which, if any, of the 5 Cs of lending should be included in investment analysis. The techniques used to perform this approach originate from very recent works in Bayesian model selection and our study is one of the first to employ these techniques in an econometric setting. Third, although several investment studies have been undertaken with an Euler equation approach, most have examined aggregate data. However, the theory behind the Euler equation approach is set at the firm level. Our study is

[2] An Euler equation is the first order condition for (the first derivative of the objective function with respect to) the variable of interest.

one of the few agricultural investment studies to use the Euler equation approach on firm level data.

Our study investigates the links between financial variables on the farm and the investment in agricultural machinery and equipment on Iowa farms. Agricultural investment in machinery and equipment is of interest because these inputs can be looked at as quasi-fixed capital. These are inputs that can be taken as fixed in the short-run, but can be varied given sufficient time and money. Investment in these inputs helps determine the long-run performance of the farm. These inputs embody technical progress and possible productivity. The value of these inputs forms a substantial portion of the farm's net worth. Additionally, the financial outlay for these inputs is quite sizable and is often not divisible.

To proceed, we outline three of the major investment models, bringing out how they were derived and what they imply. We provide a brief literature review of past studies of agricultural investment and other studies which outline our approach to the investment issue in Chapter 2. In Chapter 3, we briefly describe the Bayesian approach and discuss model selection and outlier detection issues. The data set is discussed in Chapter 4 along with a detailed description of the Bayesian composite regression model and computational strategy. The results from the Bayesian model are given in Chapter 5. In Chapter 6, we outline the Euler equation approach and estimation techniques. Chapter 7 provides the results from that analysis and contains a discussion on the Euler equation approach and the factors which may have contributed to the results. In Chapter 8, implications and conclusions are drawn.

CHAPTER 2. REVIEW OF ECONOMIC LITERATURE

2.1 Investment Models[1]

There are three main approaches to modeling investment behavior: the accelerator, neoclassical, and Q models of investment. In the accelerator model of investment (Clark, 1917), investment is seen as strictly a function of a change in output. The simple accelerator model is

$$I_t = \alpha \Delta Q_t + \varepsilon_t$$

where I_t, net investment, is equal to $K_t - K_{t-1}$, K_t is the capital stock, Q_t is the output level, Δ represents a difference operator, ε_t is a stochastic error term, and α is a unknown parameter to be estimated. The model assumes there are no delivery lags for the capital purchased; the age of the existing capital (vintage effect) is immaterial; and there are no adjustment costs arising from the addition of new capital to the production process.

Several studies found the simple accelerator model to be insufficient (Tinbergen, 1938; Chenery, 1952; Koyck, 1954). These results led to the model being expanded to allow for delivery lags or for expectations of future output to be based upon previous output changes. This expanded model is the flexible accelerator model and the investment equation for it is

$$I_t = \Sigma_{j=0}^{J} \alpha \beta_j \Delta Q_{t-j} + \varepsilon_t$$

[1] Most of this section is based upon an excellent review of the investment modeling literature provided by Chirinko (1993).

where the β_js represent lag parameters for J lags. In both of these models, (input and output) price effects are excluded. It is also assumed that the firm has enough access to funds to meet any investment needs and that the way in which the funds are obtained does not affect the investment process.

In response to some of the theoretical lapses of the accelerator models, Jorgenson (1963, 1971) and others proposed the neoclassical model. Under the neoclassical model, the firm is assumed to maximize its discounted profit stream over an infinite horizon. Capital depreciates at a geometric rate. There are no delivery lags, adjustment costs, or vintage effects. The optimal capital level is determined by output and the user cost (rental price) of capital. The firm can reach its optimal capital level immediately. However, the investment relationship that is normally estimated under the neoclassical model assumes delivery lags for new capital. Assuming the production function holds a constant elasticity of substitution (σ) between variable inputs and capital, the investment equation for the neoclassical model is given by

$$I_t = \delta K_{t-1} + \Sigma_{j=0}^{J} \alpha \beta_j \Delta\left(Q_{t-j}\left(C_{t-j}{}^{-\sigma}\right)\right) + \varepsilon_t$$

where C_t is the user cost (rental price) of capital[2] and δ is the depreciation rate. If σ is zero, then the neoclassical model reduces to the flexible accelerator model. If, in addition, the new capital delivery lags are removed, then the neoclassical model becomes the simple accelerator model.

---

[2] This variable is be described in greater detail in Section 4.3.

However, several problems exist with the neoclassical model. For instance, the simultaneity of the firm's choices of output and capital stock is not adequately addressed. Also, the optimal capital stock assumes no delivery lags, while the investment equation takes delivery lags into account. The neoclassical model does open up investment behavior to the effects of price through the user cost of capital; however, there is no accounting for the sources of the investment funds or their effect on the investment decision.

The Q theory of investment more formally addresses expectations involved in the investment process. First introduced by Keynes in 1936 and reintroduced by Brainard and Tobin (1968) and Tobin (1969), formal models under this theory have linked investment with marginal Q, the ratio of the discounted future revenues from an additional piece of capital over the purchase price of the capital. In this model, the firm is assumed to maximize its market value (the discounted sum of expected profits). The firm is a price-taker in all markets (input and output) and faces adjustment costs when it deviates from the "average" or "normal" investment rate. These adjustment costs are assumed to be convex which forces the firms to consider their expectations about the future.

The financial value of firms whose stock is traded on organized exchanges can be easily ascertained. Expectations about the future financial value of the firm are embedded in the stock price. Thus, average Q (the ratio of the financial value of the firm over the replacement cost of its existing capital stock) is observable, but marginal Q is not due to its dependence upon expectations of future revenues from the additional piece of capital. For estimation purposes, average Q replaces marginal Q in the investment equation for the Q model. The typical investment equation used for estimation of the Q model is

$$\frac{I_t}{K_t} = \left(\frac{1}{\alpha}\right)\left(Q^A{}_t - 1\right)p^I{}_t + \varepsilon_t$$

where $Q^A{}_t$ represents average Q and p^I_t is the relative price of investment (relative to output).

Marginal Q is equal to average Q only under certain conditions (Hayashi, 1982, 1985). These are:

1) input and output markets are competitive,
2) the production and adjustment cost functions are linearly homogeneous,
3) capital is homogeneous, and
4) investment decisions are separate from other financial decisions.

Thus, to apply this model to investment, one again assumes that funds are readily available and that their source is immaterial to the investment decision. Another problem in applying this model to agriculture is that most farms do not sell stocks in financial markets, making Q hard to determine.

In each of these models, internal and external finance are treated as perfect substitutes. The firm is unconcerned or unaffected by the choice of internal or external funds. This would be true if financial capital markets were perfect. There would be no transaction costs or asymmetric information problems between lenders and borrowers. Although this type of assumption may be adequate in some studies, it is hard to justify for agricultural investment at the farm level. Many studies have investigated agricultural lenders' credit rating procedures, essentially examining the asymmetric information problem. These studies suggest that asymmetric information problems are significant in agricultural lending. With this study, we seek to find the connection between the asymmetric information literature and the investment literature by including financial variables into investment models. In the Bayesian composite regression approach, elements from the investment models above are combined with variables representing the 5 Cs of the farmer credit situation and the absolute and relative impacts of these variables on farm machinery investment are explored. In the Euler equation approach, a borrowing constraint, also incorporating the 5 Cs, is inserted into the standard model to examine whether financial variables can help explain investment decisions.

In the following section, we review the recent literature of investment studies, categorizing them by their sector of interest. Because most of the innovations in the investment literature have originated in the study of non-agricultural investment, this literature is reviewed first. Adaptations of these techniques to agricultural investment are then covered. We conclude with a brief look at agricultural credit studies.

2.2 Non-agricultural Investment Studies

Most of the investment studies within the past two decades can be placed into three classes: variations on the Q model, Euler equation studies, or eclectic studies which combine components from several different models such as those discussed above. Fazzari and Mott (1986) is an example of an eclectic approach to investment modeling. In their paper, they modeled investment as a function of sales, internal finance, and acceleration variables. They proposed a positive relationship between investment and capital utilization, proxied here by lagged sales. They also proposed a positive relationship between investment and internal finance (the sum of after-tax profits and depreciation allowances minus dividends) and a negative relationship between investment and payment commitments, represented by lagged interest expenses. Looking at data for U. S. manufacturing firms from 1967 to 1982, they found support for all three of their propositions.

Fazzari and Athey (1987) continued along this vein. They examined the investment patterns of 637 manufacturing firms from 1975 to 1985. Their investment model combined elements from the accelerator and neoclassical models with internal financial and short-term payment variables. The results indicated that the internal financial and short-term payment variables have significant impacts on investment not already covered by the standard models.

The effects of working capital on investment were explored by Fazzari and Petersen (1993). Working capital is defined as current assets (cash, inventories, accounts receivable) minus current liabilities (short-term debt and accounts payable). The authors studied investment of U. S. manufacturing firms over 1970 to 1979. Their model appended variables for the firm's cash flow and working capital to a typical Q model equation. Under their theory, investment and working capital are competing for the firm's funds and, thus, would have a negative relationship. Their findings supported this. In addition, they also expanded their equations to include sales, the change in long-term debt, and lagged investment. Their results remained robust to these changes.

Fazzari, Hubbard, and Petersen (1988) explored the impacts of financial variables on investment and examined whether these impacts vary by type of firm. Their models represented various combinations of Q and accelerator models with additional cash flow variables. Firms were classified by their dividend policy. The authors found that cash flow has a significant impact on investment, and that impact is greater for low-dividend firms.

Under both internal funds and Q models of investment, Chirinko and Schaller (1995) explored why firm liquidity would matter in investment. They divided a panel of 212 Canadian firms by three criteria: maturity, owner concentration, and group membership. It was thought that firms that are younger, that have many owners, or that do not belong to an industrial group would experience asymmetric information problems with prospective lenders. Thus, these firms might face financing constraints which would create a cost difference between internal and external funds. In the internal funds models, investment was taken to be a function of sales and internal funds. The results indicated that internal funds alleviate short-term financial constraints, therefore affecting investment timing, but did not change the optimal capital stock. Also, the firms considered more likely to face financial constraints showed larger impacts from internal funds variables in their investment decisions. The Q model specifications were augmented with firm liquidity variables. Under the Q models, the authors found that liquidity does matter in investment.

Gilchrist and Himmelberg (1995) employed a vector autoregression approach to construct a better proxy for marginal Q than average Q.[3] They labeled their proxy fundamental Q because it was based on observable "fundamentals" for the expected value of marginal Q. To separate the effects on investment of cash flow as a "fundamental" for marginal Q and as an indicator of capital market imperfections, two vector autoregressions were estimated, with and without cash flow. They estimated the standard Q equation with both average Q and fundamental Q, respectively, for the full sample of manufacturing firms and for several subsamples based on the possibility of financial constraints. Their results

---

[3] Often referred to in the literature as Tobin's Q.

showed greater support for the model with fundamental Q versus average Q. Also, the unconstrained subsamples provided stronger support for the model than the constrained subsamples. To explore the role of cash flow in greater detail, they appended cash flow to the Q equations. For the constrained firms, cash flow had a significant impact on investment.

Chirinko and Schaller (1996) examined the impacts of "bubbles" on investment and how the Q and Euler equation models of investment responded under such bubbles. A bubble is defined as the situation where the stock price for a firm deviates from the expected value of its future cash flow.[4] They examined annual data for the U. S. nonfinancial corporate sector from 1911 to 1987. Given their hypotheses, their tests distinguished among three cases. Both the Q and Euler equation model would be supported if there were no bubbles in the stock market. The Euler equation model would be supported, but the Q model would not be, if there were bubbles in the stock market, but they did not affect the investment decision. Both models would fail if the bubbles affected investment. Their results suggested that bubbles exist, but they do not affect investment.

Bond and Meghir (1994) investigated investment sensitivity to internal funds. To include internal financial variables theoretically into their model, the authors employed a hierarchy of finance approach. Under this approach, internal funds are cheaper than external funds (debt and equity issues). This departed from most investment models which assume there is no difference between internal and external funds, except possibly for tax considerations. They estimated a regression based upon the investment Euler equation using

[4] The fundamental value of a firm's stock is the expected present value of its future cash flow.

GMM for a panel data set of British firms. They found evidence that the firm's liquidity matters in investment decisions.

Using an Euler equation specification, Hubbard, Kashyap, and Whited (1995) studied the investment patterns of manufacturing firms from 1976 to 1987. They added a borrowing constraint to the standard model to incorporate the effects of cash flow and overall macroeconomic borrowing conditions. Their analyses were performed on the full sample and on subsamples based on the size and dividend of the firms. For the standard model without the borrowing constraint, the results supported the model for high dividend firms, but not for low dividend firms. Both small and large firm subsamples rejected the standard model. However, when the model was augmented with the borrowing constraint parameterized with cash flow and a measure of national borrowing conditions, the results from all subsamples supported the model. This indicated that for financially constrained firms, both the internal and national financial situations had an impact on the firm's investment decision.

In an earlier paper, Whited (1992) employed an Euler equation approach to investigate the relationship between debt, liquidity constraints, and investment for 325 U. S. manufacturing firms. She estimated both the traditional model and an augmented model with a borrowing constraint. The borrowing constraint specification provided an avenue for the influence of firm financial variables. The two financial variables that were chosen for the model are the firm's debt to asset ratio and interest coverage ratio. The interest coverage ratio is the ratio of interest expenses to the sum of interest expenses and cash flow. The debt to asset ratio could be interpreted as a measure of the firm's collateral or the firm's current relative capacity for debt. The interest coverage ratio could be interpreted as a measure of financial distress. As other studies have done, the models were estimated on the full sample and various subsamples. The subsamples were based on the levels of the debt to assets and interest coverage ratios, and bond ratings. The results provided evidence that the inclusion of the borrowing constraint improves the model and thus pointed to a role for financial variables in the investment decision. The sample splits indicated that as firms are considered to be more financially constrained, the more likely the traditional model is rejected in favor of the augmented model with the borrowing constraint (i.e., the more likely firm financial variables have an impact on investment).

2.3 Agricultural Investment Studies

Gustafson, Barry, and Sonka (1988) used an experimental simulation approach to study agricultural investment. In their study, farmers were presented with four policy scenarios under which they would make investment decisions for their farm. Each scenario was repeated to create a four year study period. The scenarios included a baseline run, a lower commodity price support run, a revision of the tax code run, and an interest rate buydown program run. Before the experiment, farmers completed surveys which provided information on their past business performance, personal and farm characteristics, expectations of the future farming situation, and a ranking of factors in their investment decision. During the experiment, expected commodity prices, yields, interest rates, and inflation rates were elicited from the farmers. Then based on these, the farmers made their investment decisions about land and machinery with no set limitations. The financial impacts of the investment were analyzed and reviewed with the farmer, at which time they could adjust their investment plan. Once the farmer had finalized the investment strategy, "actual" prices, yields, interest rates, and inflation rates were revealed and the "actual" financial standing of the farm was computed. This process was repeated to achieve the four year study period. From the investment simulation results, the authors found that several financial variables, such as farm leverage (the ratio of the farm's total liabilities and net worth), were significant in influencing investment.

Weersink and Tauer (1989) constructed traditional and dynamic investment models based upon the flexible accelerator model. To the traditional model, they also incorporated alternative investment models by including measures of the cost of capital, profit expectations, desired capital stock, real liability, real farm net income, farm size, operator age, and a time trend. Thus, their version of the traditional investment model was a composite regression approach, combining elements from several different investment models. Using a panel data set of 112 dairy farms over 10 years (1974-83) from the New York Dairy Farm Business Summary, they found that the traditional investment model performed better that the dynamic model. Several variables were found to be significant in the investment decision, including net farm income and liabilities.

In their study of farm business expansion, LaDue, Miller, and Kwiatkowski (1991) estimated the probabilities of expansion based on three categories: no investment, replacement investment, and expansion (investment above replacement). They considered eleven independent variables in the study: farm size (as measured by gross income), operator age, equity ratio (net worth / total assets), farm goals, education, a management index, geographic location, the interest rate, urban proximity, income expectations, and farm type. They estimated ordered logit models for a sample of New York farms in 1985 and 1986. Only two of the variables were consistently significant in the farmer's investment decision, operator age and gross income. They found that large farms and young farmers were the most likely to expand, while older farmers and small farms were likely to make no investment.

Jensen, Lawson, and Langemeier (1993) built a composite model based upon accelerator and neoclassical investment models and added internal cash flow variables. The cash flow variables were justified by pointing to the studies of agricultural lenders' methods for evaluating a farm's credit rating. A linear regression composed of the important variables from each of these prospectives was estimated for a panel data set of 522 farms over 16 years (1973-88) from the Kansas Farm Management Association (KFMA). The results indicated that variables from all three categories are important. The elasticity measures of investment were most responsive to the cash flow variables.

Following the lead of Weersink and Tauer, Chellappan and Pederson (1995) formed an agricultural machinery investment model that included elements from the accelerator, neoclassical, and internal funds models. They also included the farmer's age, machinery age, and total liabilities as other explanatory variables. Their data set was an unbalanced panel data set of 116 farms over 6 years (1985-90) from the Minnesota Farm Business Management Association. Farms were restricted to have at least 70 percent of farm revenues from crop sales. The authors estimated two forms of their model: a two-way fixed effects accounting for individuals and years, and a random effects model for individuals. The results of the two forms were nearly the same. Output, profit expectations, lagged capital, machinery age, and total liabilities were significant in their linear regressions.

Bierlen and Featherstone (1998) employed a fundamental Q model approach to agricultural investment. Fundamental Q, as discussed in Gilchrist and Himmelberg (1995), is a measure of the expected discounted marginal profit stream from an additional dollar invested. The authors estimated the fundamental Qs for 405 KFMA farms over the 1973-95 period. Investment was then taken to depend upon fundamental Q and cash flow. The model was estimated over the full sample and for selected subsamples based on operator age, farm assets, debt-to-asset ratios, and time period. Both fundamental Q and cash flow were found to be significant factors in agricultural investment. The time period subsample results indicated that credit constraints were not a significant problem in the 1970s, but the financial markets became tighter during the 1980s and early 1990s.

Using an Euler equation approach, Hubbard and Kashyap (1992) examined aggregate U.S. agricultural investment in equipment. They incorporated a financial constraint into the model by assuming that outstanding debt is less than a debt ceiling. This financial constraint was assumed to hold in periods when collateralizable net worth is low. A generalized methods of moments (GMM) estimation technique was employed in order to estimate the nonlinear model while taking simultaneity problems into account. The authors found support for the role of internal funds variables in investment models. One of the main critiques of this paper is the use of aggregate data, when the Euler equation model is based on firm level theory.

Bierlen and Featherstone (1996) also applied an Euler equation approach to agricultural investment. Models were structured assuming either expected profit or expected utility maximization; thus, the study examined both the risk neutral and risk averse cases. As in Hubbard and Kashyap, a debt constraint (based on net worth and a risk index) was added to the standard model. The farmer's utility function was taken to be negative exponential and the parameter estimates were arrived at through GMM. Farm level data for 397 KFMA farms over 1975 to 1992 were used in the analysis, thus avoiding the critique faced by the Hubbard and Kashyap study. The models were estimated over the entire data set and data splits for farm size, debt-to-asset ratio, and farmer age. The debt constraint was found to be significant, thus implying that the farms face credit rationing. The constraint affected small farms, high debt farms, and older farmers the most.

2.4 Agricultural Credit Studies

Many of the studies above found financial variables to be significant in the investment decision. Hence, we have reviewed the agricultural credit literature to find the factors that most influence the credit decision from a lender's point of view. Rather than give an extensive listing of the studies in this area, we have chosen two papers which summarize the existing literature and help form the basis for our variable choices.

In their study of credit assessment models, Miller and LaDue (1989) summarized the results from nine other agricultural credit studies. Within these nine studies, 23 different factors had been shown to be important in assessing borrower quality. Of these factors, only measures of solvency (owner equity), repayment ability, or liquidity appeared significant in a majority of studies. From this summary, the authors formed a model for the probability of loan default. Since many measures can be constructed to represent farm financial characteristics, an analysis of variance was performed to select the independent variables for the model. Variables representing liquidity, solvency, profitability, and operating efficiency were selected for the model. Results from fitting a logistic regression model indicated that the liquidity, profitability, and operating efficiency variables were significant in assessing borrower quality.

Knopf and Schoney (1993) looked at the use of several economic, efficiency, and financial variables to explain agricultural loan success rates. They outlined what banks and other lending agents seek in loan clients. These attributes were referred to as the "5 Cs": character, capacity (cash flow), collateral, credit rating, and capital (owner's equity). They began with a list of 59 candidate variables for their logit regression. To select which variables to include in their final model, the authors employed a forward selection technique based on Wald statistics.[5] Twenty of the variables were chosen for the model. The final results showed that most traditional financial ratios added little to the loan success rate, only the current ratio, the ratio of short-term assets to short-term liabilities, was significant. The authors outlined several reasons why this result may have occurred, the main reason being inconsistent variable definition and measurement.

We have presented a review of the literature on investment and the factors that may be associated to it. In the next chapter, we take a step back and review the literature on the statistical methods that are used to identify these relationships.

---

[5] Model selection techniques are discussed in more detail in Section 3.3.

CHAPTER 3. REVIEW OF STATISTICAL METHODS AND LITERATURE

3.1 The Bayesian Approach

When performing research, we often draw conclusions about phenomena based on observed data. The techniques used to analyze and summarize data vary, depending on the questions asked and the perceptions of what is required to answer the question. In most econometric work, the "answers" to the questions are summarized in point parameter estimates and confidence intervals derived from classical statistical analyses. Under a Bayesian framework, however, results of analyses are summarized into probability distributions. We now give a brief introduction to the Bayesian paradigm. For a more complete discussion, see Box and Tiao (1973) and Gelman et al. (1995), for example.

Let $(\theta_1, \theta_2) \in \Theta$ represent two scalar-valued parameters, and let y denote a vector of observations. In the Bayesian approach, all parameters are considered to be random variables. The goal in most Bayesian analyses is to estimate the distribution of (θ_1, θ_2) using information provided by the data, in addition to any prior information about (θ_1, θ_2) that might be available.

To proceed, we use Bayes' theorem (or rule), which states

$$p(\theta_1,\theta_2 \mid y) = \frac{p(y \mid \theta_1,\theta_2)p(\theta_1,\theta_2)}{\int_\Theta p(y \mid \theta_1,\theta_2)p(\theta_1,\theta_2)\partial\theta_1\partial\theta_2},$$

where $p(\theta_1, \theta_2)$ is the prior distribution of (θ_1, θ_2), $p(y \mid \theta_1, \theta_2)$ is the usual likelihood function, and

$$p(\mathbf{y}) = \int_\Theta p(\mathbf{y} \mid \theta_1, \theta_2) p(\theta_1, \theta_2) \partial\theta_1 \partial\theta_2$$

is the marginal distribution of the data, or normalizing constant. Often, the normalizing constant is omitted, and Bayes' rule is written in its "proportional" form

$$p(\theta_1, \theta_2 \mid \mathbf{y}) \propto p(\mathbf{y} \mid \theta_1, \theta_2)\, p(\theta_1, \theta_2).$$

Inferences about, e. g., θ_1, are based on the marginal distribution of θ_1,

$$p(\theta_1 \mid \mathbf{y}) = \int p(\theta_1, \theta_2 \mid \mathbf{y}) \partial\theta_2\,,$$

obtained by integrating the joint posterior distribution $p(\theta_1, \theta_2 \mid \mathbf{y})$ with respect to the "nuisance" parameter θ_2. In general, the parameter vector of interest has dimension k, possibly large, and thus calculation of the normalizing constant and of the various marginal densities is difficult. We discuss this issue in the next section.

To proceed as a Bayesian, we must first construct a full probability model for all observed and unobserved quantities in our problem. The model is written as a joint distribution of data and parameters, $p(\mathbf{y}, \theta)$, and can be decomposed into two pieces, the sampling distribution or likelihood (the conditional distribution of the data given the parameters) and the prior distribution (the marginal distribution of the parameters). The prior distribution represents all of the information available about the parameters before the analysis is conducted. The likelihood function reflects information about θ that is provided by the data $\mathbf{y}$. An important objective of our analysis is to calculate the posterior distribution, $p(\theta \mid \mathbf{y})$, the conditional distribution of the parameters given the data. The posterior distribution represents the updated information about the parameters available after combining prior and sample information. Choosing prior distributions for parameters is not trivial. Berger (1985) and Bernardo and Smith (1994) give good discussions on the topic. Briefly, we distinguish between informative and noninformative prior distributions, and between proper and improper prior distributions. In this paper, we use both informative and diffuse priors, but limit ourselves to distributions that are integrable.

In many problems, the Bayesian approach presents advantages over the frequentist viewpoint. When the analysis involves several steps, the Bayesian framework permits accounting for uncertainties about parameters that are accumulated along the way. Credible intervals, the Bayesian equivalent to frequentist confidence intervals, have a more appealing interpretation for practitioners. Modern numerical methods provide a simple mechanism for estimating posterior distributions of any (measurable) function of the parameters in the model. As a result of recent advances in computing and of the many numerical approaches now available to practitioners, the Bayesian framework is used these days to fit highly complex models to large data sets.

3.2 Numerical Procedures[1]

As was discussed above, applying Bayesian methods requires integration, often in many dimensions. In our problem, for example, we would need to integrate the joint density of the data and parameters in over 500 dimensions. Except in the few cases in which analytical integration is possible, or in trivial problems involving just a few dimensions, applying Bayesian methods was all but impossible until recently. In 1990, Gelfand and Smith

[1] This section is derived from Chapter 11 in *Bayesian Data Analysis* by Gelman et al. (1995) and Brooks (1998).

introduced to statisticians the Gibbs sampler (first proposed by Geman and Geman, 1984) making it possible for practitioners to apply Bayesian methods to realistic, complex problems. The Gibbs sampler is one of a family of algorithms called Markov chain Monte Carlo (MCMC) methods.

The literature leading to current MCMC methods can be traced backed to Metropolis and Ulam (1949) and Metropolis et al. (1953). These papers constructed a tool for Markov chain simulations of probability distributions, the "Metropolis algorithm." Hastings (1970) extended these results and indicated the potential for applications in statistical analysis. In their study of image restoration, Geman and Geman (1984) proposed what has become known as the Gibbs sampler. Gelfand and Smith (1990) brought the Gibbs sampler to the attention of mainstream statistical research.

We now give a brief description of MCMC methods for approximating marginal posterior distributions. Suppose we are interested in the distribution of the parameter vector θ and have data y. MCMC methods are employed when the posterior distribution cannot be obtained in closed form. For example, $p(\theta \mid y)$ cannot be obtained in closed form if we cannot compute the normalizing constant $p(y)$. The idea behind MCMC techniques is simple: generate draws, θ^t $(t = 1, 2, \ldots)$, from the distribution of interest $p(\theta \mid y)$ by generating a Markov chain in θ whose stationary distribution is equal to $p(\theta \mid y)$. We use the term "target distribution" to refer to the distribution of interest, $p(\theta \mid y)$.

An MCMC simulation proceeds in the following way. We choose θ^0, a starting point for the chain.[2] Then, for each iteration t (t = 1, 2, ...), we draw θ^t from a transition distribution $p(\theta^t \mid \theta^{t-1})$ where the transition distribution is constructed so that the Markov chain will converge to the target distribution. After many iterations, the simulated values from the chain can be considered as a (dependent) sample from the distribution of interest and can be used to obtain summary statements about the target distribution. The Gibbs sampler[3] is a particular form of an MCMC simulation. In the Gibbs sampler, the Markov chain in θ is constructed by drawing values of θ from its conditional distribution, given the value of θ in the previous step. For example, let $\theta = [\theta_1, \theta_2, \ldots, \theta_m]'$ and at iteration t let $\theta_j^t \sim p(\theta_j^t \mid \theta_{-j}^{t-1}, y)$ where $\theta_{-j}^{t-1} = [\theta_1^t, \ldots, \theta_{j-1}^t, \theta_{j+1}^{t-1}, \ldots, \theta_m^{t-1}]'$. It has been shown (e.g., Besag, 1974) that the Markov chain formed by draws $\theta_j^1, \theta_j^2, \ldots$ has a stationary distribution equal to $p(\theta \mid y)$ (j = 1, 2, ... , m), as long as certain conditions hold. The Gibbs sampler is particularly convenient when the conditional distributions are of standard forms (such as normal, gamma, etc.).

For MCMC methods to work, the chains must have a unique stationary distribution, the target distribution. If the Markov chains are irreducible, aperiodic, and non-transient, then they will have a unique stationary distribution.[4] The irreducibility requirement means that any point in the parameter space can be reached from any other point in the space. Periodicity refers to the probability of returning to a given state. A Markov chain is periodic

[2] Superscripts denote iterations or links in the chain.
[3] Also referred to as alternative conditional sampling.
[4] The definitions of the Markov chain properties are derived from Feller (1968) and Gelman et al. (1995).

with period d if the n-step transition probability[5] $p^n(\theta^* \mid \theta^*) = 0$ unless $n = md$ for m an integer. A Markov chain is aperiodic if $d = 1$. The non-transiency property states that the waiting time for the chain to return to a state is finite. Given these conditions, the chain will have a unique stationary distribution and that distribution will be the target distribution. Geman and Geman (1984) showed that the Gibbs sampler satisfies convergence, rate, and ergodicity properties. The convergence property states that the joint distribution $[\theta_1^t, \theta_2^t, \ldots, \theta_m^t]$ converges in distribution to ($\xrightarrow{d}$) $[\theta_1, \theta_2, \ldots, \theta_m]$ and hence that $\lim_{t\to\infty} \theta_j^t \xrightarrow{d} \theta_j \sim p(\theta_j)$ for all j.[6] The rate of convergence of the joint distribution is geometric in iteration time t. Ergodicity states that for any measurable function $f(\theta_1, \theta_2, \ldots, \theta_m)$ whose expectation exists, $\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} f(\theta_1^t, \theta_2^t, \ldots, \theta_m^t)$ converges almost surely to $E[f(\theta_1, \theta_2, \ldots, \theta_m)]$.

Many methods will produce chains which satisfy the requirements specified above. The differences among these methods is in the definition of the transition distribution and probability. Following Gelman et al. (1995), we show the transition distributions and probabilities for the Metropolis-Hastings algorithm. Suppose we are interested in the target distribution $p(\theta \mid y)$. Let $T_t(\theta_a \mid \theta_b)$ represent a transition distribution, $J_t(\theta_a \mid \theta_b)$ represent a jumping distribution, and r represent the transition probability. For each iteration, algorithms proceed by drawing θ^t from $T_t(\theta^t \mid \theta^{t-1})$. Given θ^{t-1}, $T_t(\theta^t \mid \theta^{t-1})$ is a mixture of the point mass

[5] The probability of moving from θ^* back to θ^* in exactly n iterations.
[6] Casella and George (1992) provided a nice convergence proof for the case of a 2 X 2 table with multinomial sampling.

$\theta^t = \theta^{t-1}$ and $J_t(\theta^t \mid \theta^{t-1})$. The transition probability (r) is determined by a ratio of importance ratios. Let θ^* be a candidate drawn from $J_t(\theta^* \mid \theta^{t-1})$. Then

$$r = \min\left(1, \frac{p(\theta^* \mid y) / J_t(\theta^* \mid \theta^{t-1})}{p(\theta^{t-1} \mid y) / J_t(\theta^{t-1} \mid \theta^*)}\right),$$

and θ^t is set equal to θ^* with probability r, or remains at θ^{t-1} with probability 1−r. For convenience and efficiency of the algorithm, there are several properties that the jumping distribution should have. The jumping distribution $J_t(\theta^t \mid \theta^{t-1})$ should be easy to sample from, the transition probability should be easy to compute, each jump should be of a reasonable size to expedite the iteration process, and rejection of jumps should be limited, so that the chain does not get "stuck."

The Metropolis algorithm and the Gibbs sampler are special cases of the Metropolis-Hastings algorithm. For the Metropolis algorithm, the jumping distribution is symmetric ($J_t(\theta_a \mid \theta_b) = J_t(\theta_b \mid \theta_a)$ for all θ_a, θ_b, and t) and r reduces to $\min\left(1, \frac{p(\theta^* \mid y)}{p(\theta^{t-1} \mid y)}\right)$. For the Gibbs sampler, we define the distributions for the j^{th} subvector and the t^{th} iteration as follows.

$$J_{j,t}(\theta^* \mid \theta^{t-1}) = \begin{cases} p(\theta_j^* \mid \theta_{-j}^{t-1}, y) & \text{if } \theta_{-j}^* = \theta_{-j}^{t-1} \\ 0 & \text{otherwise} \end{cases}$$

and r equals one. Thus, in the Gibbs sampler, every jump is accepted.

For a (somewhat trivial) example[7], consider the standard regression model: $\mathbf{y} \sim N(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$, where $\mathbf{y}$ is an n x 1 vector of observations, $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_k]$ is the n x k matrix of regressors, $\boldsymbol{\beta} = [\beta_1, \beta_2, \ldots, \beta_k]'$ is a k x 1 vector of parameters, σ^2 is a scalar, and $\mathbf{I}$ is the n x n identity matrix. Here, $\theta = (\beta, \sigma^2)$. The standard noninformative prior distribution for (β, σ^2) is $p(\beta, \sigma^2) \propto \sigma^{-2}$. For this model, the conditional marginal distributions for β and σ^2 are given by $p(\beta \mid \sigma^2, \mathbf{y}) \sim N(\hat{\beta}, \sigma^2(\mathbf{X}'\mathbf{X})^{-1})$ and $p(\sigma^2 \mid \beta, \mathbf{y}) \sim$ Inverse $\chi^2(n-k, s^2)$ where $\hat{\beta}$ and s^2 are the classical least squares estimates for β and σ^2. Gibbs sampling could proceed using $\hat{\beta}$ as a starting value. Draw $(\sigma^2)^1$ from $p(\sigma^2 \mid \hat{\beta}, \mathbf{y})$ and draw β^1 from $p(\beta \mid (\sigma^2)^1, \mathbf{y})$ to complete the first iteration of the sampler. Repeat these two steps a large enough number of times so that the chain "converges."

Several issues are of concern with iterative simulations such as the Gibbs sampler. For example, the effect of starting values, the dependence in the chains, and the criterion for convergence of the sequences may significantly affect results. Gelman et al. (1995) recommended that several simple procedures be followed to alleviate these concerns. Multiple chains, rather than just one, can be generated for each parameter to reduce the dependence among sample elements, and to permit computation of simple convergence diagnostics. The impact of the starting values can be reduced by discarding the first portion of draws in analyses. The behavior of each scalar estimand can be monitored to decide when the chains have converged.

[7] We use this problem only to illustrate the Gibbs sampler in a simple, but widely used framework.

The initialization of the chains can be done in a variety of ways. The idea is to guarantee that any value in the parameter space has positive probability of being chosen as an initial value for the chain. One technique consists in selecting starting values as draws from an overdispersed distribution. For example, in the regression example above, we could select starting values for β from random draws from $N(\hat{\beta}, s^2(X'X)^{-1})$ or from a *t*-distribution with the same mean and scaled to have a variance at least as large as $s^2(X'X)^{-1}$. Using multiple chains initialized in this way permits exploring the parameter space quickly and monitoring the behavior of chains for determining convergence.

The behavior of chains is monitored by Gelman and Rubin's R-statistic, $\sqrt{\hat{R}}$. The R-statistic computes the potential scale reduction in the current distribution of the scalar estimand if the simulations were continued to infinity. The statistic compares the relative sizes of the between- and within-sequence variances, and provides an intuitively appealing stopping point when both variances are about equal. Let λ be a scalar and suppose we have generated J chains (sequences) with N draws each (after removing the first halves of the chains) from a Gibbs sampler run. Let λ_{ij} represent the i^{th} draw from the j^{th} chain ($i = 1, 2, \ldots, N; j = 1, 2, \ldots, J$). The between-sequence variance is given by:

$B = \frac{N}{J-1}\Sigma_{j=1}^{J}(\bar{\lambda}_{.j} - \bar{\lambda}_{\_})^2$ where $\bar{\lambda}_{.j} = \frac{1}{N}\Sigma_{i=1}^{N}\lambda_{ij}$ for each j and $\bar{\lambda}_{\_} = \frac{1}{J}\Sigma_{j=1}^{J}\bar{\lambda}_{.j}$. The within-sequence variance is given by: $W = \frac{1}{J}\Sigma_{j=1}^{J}s_j^2$ where $s_j^2 = \frac{1}{N-1}\Sigma_{i=1}^{N}(\lambda_{ij} - \bar{\lambda}_{.j})^2$. Gelman and Rubin's R-statistic is computed as $\sqrt{\hat{R}} = \sqrt{\frac{V}{W}}$ where $V = \frac{N-1}{N}W + \frac{1}{N}B$. Using the fact that $\lim_{N \to \infty} \sqrt{\hat{R}} \to 1$, convergence is considered to be achieved when all scalar estimands have R-statistics near one.[8]

3.3 Model Selection Techniques

Model selection refers to the search for subsets of covariates that best associate to a dependent variable given a decision rule. Many techniques have been proposed from both the frequentist and the Bayesian viewpoints. Here, we outline several of these techniques.[9] Frequentist methods are based on statistics such as the coefficient of determination (R^2), the adjusted coefficient of determination ($\overline{R}^2$), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC). Automatic selection methods such as stepwise, forward, and backward selection algorithms have also been proposed from a frequentist perspective. Bayesian techniques include model averaging and Stochastic Search Variable Selection (SSVS). We now briefly describe each of these techniques.

Consider the following linear model, $\mathbf{y} = \mathbf{X}\beta + \varepsilon$, $E[\varepsilon \mid \mathbf{X}] = 0$, $\mathrm{Var}(\varepsilon \mid \mathbf{X}) = \sigma^2\mathbf{I}$, where $\mathbf{y}$ is an n x 1 vector of observations, $\mathbf{X}$ is an n x k matrix of covariates with the first column being a vector of ones, β is a k x 1 vector of unknown parameters, ε is an n x 1 error vector, and $\mathbf{I}$ is the n x n identity matrix. The variance component, σ^2, is unknown. With linear models such as this, one of the most commonly used evaluation criteria is the coefficient of determination, R^2. The coefficient of determination is defined as one minus the

[8] Gelman et al. (1995) suggested values below 1.2 are acceptable. However, target levels for the R-statistics should be set according to the level of precision required.

[9] For a more detailed discussion on the classical techniques, see Gourieroux and Monfort (1995), Grasa (1989), Miller (1990), or Snedecor and Cochran (1989).

ratio of the residual to the total sum of squares. It measures the proportion of the variation in the dependent variable that can be accounted for by the model. Models with higher R^2 provide better predictions over the sample space than those with lower R^2. However, the coefficient of determination has a significant drawback as a model selection tool. It increases with the number of covariates in the model; thus, this measure always leads to larger models.

The adjusted coefficient of determination, $\overline{R}^2$, was proposed to avoid this problem. It adjusts R^2 for degrees of freedom using the formula: $\overline{R}^2 = 1 - \frac{n-1}{n-k}(1 - R^2)$. The adjusted coefficient of determination, $\overline{R}^2$, is always equal to or less than R^2 and the difference increases with the number of explanatory variables. Thus, $\overline{R}^2$ penalizes larger models, unless the additional covariates are significantly associated to y. Model choice based on maximizing $\overline{R}^2$ is equivalent to model selection using a minimum s^2 criterion, where s^2 is the best quadratic unbiased estimator of σ^2. The adjusted coefficient of determination, however, suffers from an inconsistency problem. Let M_i and M_j be two normal linear models with M_i nested in M_j. When M_i is the "correct" model, the probability of choosing M_i using the $\overline{R}^2$ criterion does not converge to one as the number of observations goes to infinity (Gourieroux and Monfort, 1995).

There are two procedures which may be referred to as stepwise regression. One, the step-up or forward selection method, begins by computing regressions of y on each individual X_i ($i = 1, \ldots, k$). The comparison among models is based on the residual mean square and the F-statistic that tests if the coefficient associated to each selected parameter is equal to zero. At each step, the F-statistic for the selected variable must exceed the chosen boundary level or the selection process is concluded. The $\mathbf{X}_i$ with the smallest residual mean square is chosen. Suppose $\mathbf{X}_j$ is chosen. Next, all k−1 regressions pairing $\mathbf{X}_j$ and $\mathbf{X}_i$ ($i = 1,\ldots, k$, $i \neq j$) are computed and the variable yielding the greatest reduction in the error sum of squares after fitting $\mathbf{X}_j$ is chosen. This process continues until the null hypothesis is no longer rejected.

The other stepwise regression approach is the step-down or backward selection method. Again the method relies on an F-test; this time variables are omitted if their F-statistic falls below the chosen boundary level. First, the regression with all variables included is calculated. Next, k regressions are computed with each $\mathbf{X}_i$ ($i = 1,\ldots, k$) excluded in turn. The $\mathbf{X}_i$ with the smallest F-statistic is dropped if its F-statistic is below the chosen level. This process continues until no remaining variables have F-statistics below the boundary. The use of stepwise regression methods breaks down classical inference procedures. The models are built on the computed F-statistics and the chosen boundaries; thus, inferences based on these statistics are no longer appropriate (Greene, 1990).

Akaike (1973) proposed a decision rule later known as AIC. The AIC for a model is the maximum conditional log-likelihood for the model minus the number of parameters in the model:

$$\mathrm{AIC} = \sum_{i=1}^{n} \log f(y_i \mid \mathbf{X}, \hat{\beta}) - k,$$

where $\hat{\beta}$ is the maximum likelihood estimator of β. Model selection is based on finding the model with the maximum AIC.[10] The AIC penalizes larger models through the subtraction of the number of parameters. The AIC suffers from the same inconsistency as the adjusted coefficient of determination when comparing two nested normal linear models where the smaller model is the "correct" one.

Schwarz (1978) produced an alternative to the AIC derived from Bayesian arguments which avoids the inconsistencies $\overline{R^2}$ and AIC have. This decision rule has been referred to as BIC.[11] Schwarz's BIC is defined as the maximum conditional log-likelihood for the model minus the product of one-half of the number of parameters in the model and the logarithm of the number of observations:

$$\mathrm{BIC} = \sum_{i=1}^{n} \log f(y_i \mid \mathbf{X}, \hat{\beta}) - \frac{k}{2}\log(n).$$

When the number of observations exceeds eight, the penalty for including more independent variables under the BIC is larger than under the AIC.

There have been several methods proposed to carry out model selection under a Bayesian framework. One such technique is Bayesian model averaging (BMA)[12] put forth by Hoeting, Raftery, and Madigan (1996). To account for model uncertainty, the modeler averages over all possible models. Let $\mathbf{M} = (M_1, \ldots, M_K)$ represent the set of all possible models and Q represent the quantity of interest (for example, a future observation). The

---

[10] Most statistical software packages compute the AIC as –2 times this definition, hence, the chosen model has the smallest AIC.
[11] Other decision rules referred to as BIC were derived by Sawa (1978) and Chow (1981, 1983).
[12] Much of the discussion of BMA follows from Hoeting et al. (1998).

posterior distribution of Q given the data is $p(Q \mid y) = \sum_{i=1}^{K} p(Q \mid M_i, y) p(M_i \mid y)$, an average of the posterior distribution of Q under each model M_i weighted by the posterior probability of M_i. As written, performing BMA can be a daunting task due to large numbers of possible models. For k possible covariates, the number of models, K, is equal to 2^k. The authors recommend two approaches to alleviate this problem: Occam's Window and Markov chain Monte Carlo model composition (MC^3).[13]

The Occam's Window algorithm narrows the candidate model list and averages over this reduced set of models. This ad hoc approach relies on two premises. First, any model that predicts the data poorly compared to the model with the best predictions is removed from consideration. This can be stated as for any model M_j for which $\frac{\max_K \{p(M_K \mid y)\}}{p(M_j \mid y)} > C$, where C is chosen by the data analyst, is excluded. Raftery, Madigan, and Hoeting (1997) set C equal to 20. Second, if the data support a submodel more than a larger model in which the submodel is nested, the larger model is removed from consideration. Candidate models are selected by comparing the posterior odds (the ratio of posterior model probabilities) $\frac{p(M_{j+1} \mid y)}{p(M_j \mid y)}$.[14] This ratio can also be expressed as the product of the prior odds and the Bayes factor for the models. The prior odds are the ratio of prior

[13] Both approaches are discussed in Raftery, Madigan, and Hoeting (1997).
[14] Model M_{j+1} has one more covariate than model M_j.

model probabilities, $\frac{p(M_{j+1})}{p(M_j)}$. The Bayes factor is the ratio of the marginal likelihoods of the models, $\frac{p(y \mid M_{j+1})}{p(y \mid M_j)}$.

Figure 3.1 graphically displays the selection process.[15] The bounds, LL and UL, are set by the data analyst. Madigan and Raftery (1994) set LL equal to 0.05 and UL equal to 1. Raftery, Madigan, and Volinsky (1996) showed that predictive performance is improved when UL is raised to 20. If model M_j is rejected, then all of the models nested within M_j are also rejected. The authors stated that this strategy often reduces the number of models to below 25 (in fact, often to one or two models) and made it possible to average across models.

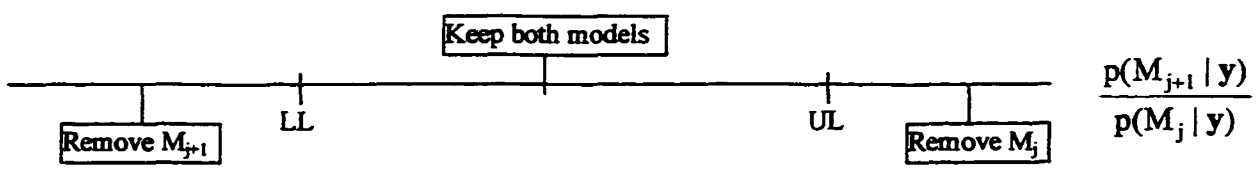


Figure 3.1. The decision rule for Occam's Window for nested models

The second approach, Markov chain Monte Carlo model composition (MC^3), approximates $p(Q \mid y) = \sum_{i=1}^{K} p(Q \mid M_i, y)p(M_i \mid y)$ through a Markov chain Monte Carlo (MCMC) approach. The Markov chain $\{M(t), t = 1, 2, \ldots\}$ is constructed with **M** being the state space and $p(M_i \mid y)$ being the equilibrium distribution. Neighborhoods, nbd(.), are

[15] The figure is based on Figure 2 of Raftery, Madigan, and Hoeting (1997).

defined for each $M_i \in \mathbf{M}$ as the model M_i and the set of models with one more or less variable than M_i. A transition matrix $\mathbf{q}$ is also defined as $\mathbf{q}(M_i \rightarrow M_j) = 0$ for all $M_j \notin \text{nbd}(M_i)$ and $\mathbf{q}(M_i \rightarrow M_j) = c$ for all $M_j \in \text{nbd}(M_i)$. Given the chain is in state M_i, we draw M_j from $\mathbf{q}(M_i \rightarrow M_j)$ and accept it with probability $\min\left\{1, \dfrac{p(M_j \mid y)}{p(M_i \mid y)}\right\}$. Otherwise the chain stays at M_i. Given this Markov chain (t = 1, 2, …, N) and certain regularity conditions, for any function $g(M_i)$ defined on $\mathbf{M}$, $\hat{G} = \frac{1}{N}\sum_{t=1}^{N} g(M(t))$ converges almost surely to $E[g(M)]$ as $N \rightarrow \infty$. Set $g(M) = p(Q \mid \mathbf{M}, \mathbf{y})$. The largest drawback to the techniques outlined for BMA is that they require either proper prior distributions or carefully constructed improper prior distributions for all parameters.[16]

George and McCulloch (1993) introduced Stochastic Search Variable Selection (SSVS). In SSVS, a hierarchical Bayesian normal mixture model describes the regression. Gibbs sampling is employed to sample from the multinomial posterior distribution and serves as a way to avoid computing all of the posterior probabilities for the numerous subsets.[17] Variables with higher posterior probability are identified as "promising" regressors. SSVS has many similarities with the MC^3 approach described earlier.

For a typical regression, $\mathbf{y} \sim N(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$, where $\mathbf{y}$ is an n x 1 vector of observations, $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_k]$ is the n x k matrix of regressors, $\boldsymbol{\beta} = [\beta_1, \beta_2, \ldots, \beta_k]'$ is a k x 1 vector of parameters, σ^2 is a scalar, and $\mathbf{I}$ is the n x n identity matrix. Not selecting $\mathbf{X}_j$ for the

[16] A prior distribution is labeled as proper if it integrates to one and does not depend on the data.

[17] Again, for a problem with k potential regressors, there are 2^k subsets of regressor combinations from which to choose.

model is equivalent to setting β_j equal to zero. The SSVS method formulates the regression model as a hierarchical model where each β_j is modeled as originating from a mixture of two normal distributions:

$$\beta_j \mid \gamma_j \sim (1-\gamma_j)N(0,\tau_j^2) + \gamma_j N(0,c_j^2\tau_j^2),$$

where γ_j equals either zero or one and $p(\gamma_j = 1) = 1 - p(\gamma_j = 0) = p_j$. The variance τ_j is set small (but greater than zero) and c_j is set large (greater than one) so that when γ_j equals one, a nonzero draw of β_j is included in the model, or when γ_j equals zero, the value of β_j would be so close to zero, it could be set at zero without significantly impacting the results. George and McCulloch (1993) provided guidelines in setting the parameters. The probability p_j represents the prior probability that X_j is included in the model.

Once prior distributions are chosen for the γ_j and σ^2, SSVS employs the Gibbs sampler to generate the following Markov chain: $\beta^0, \sigma^0, \gamma^0, \beta^1, \sigma^1, \gamma^1, \ldots$, where β^0 is the least squares estimate of β, σ^0 is the least squares estimate of σ^2, and γ^0 is a vector of ones.[18] Parameters β^t, σ^t, and γ^t ($t = 1, 2, \ldots$) are drawn from their conditional distributions. Once convergence of the sequence is attained, the γ draws provide evidence on promising regressor subsets based on posterior probabilities. George and McCulloch (1997) extended the SSVS method to the case of more general models. Criticisms of the SSVS method are that, in the original formulation, regressors are never actually removed from the model but their parameter is set close to zero with a high probability (Raftery, Madigan, and Hoeting,

[18] These settings initialize the Gibbs sampler.

1997) and the variable selection process is based on practical significance, not necessarily statistical significance.

Geweke (1996) presented an approach to variable selection similar to SSVS and MC^3. The prior distributions for the parameters are mixtures of normal distributions and point masses at zero (to indicate the variable is not selected for the model). Computation is performed using a Gibbs sampler with complete blocking. Each parameter is drawn from its distribution conditional on the values of all of the other parameters.[19] To define these conditional distributions, we look at a simplified model. For β_j, given β_p ($p = 1, 2, \ldots, k$; $p \neq j$) and σ, we define $z_i = y_i - \Sigma_{p \neq j}\beta_p X_{i,p}$. The simplified model is then $z_i = \beta_j X_{i,j} + \varepsilon_i$, $\varepsilon_i \sim N(0, \sigma^2)$ ($i = 1, \ldots, n$). The decision on whether $\beta_j = 0$ or $\beta_j \neq 0$ is based on a comparison of a draw from a uniform(0, 1) random variable and the conditional posterior probability ($\overline{p}_j$) that $\beta_j = 0$. This conditional posterior probability is calculated as $\overline{p}_j = \dfrac{\underline{p}_j}{\underline{p}_j + (1 - \underline{p}_j)BF}$ where $\underline{p}_j$ is the prior probability that $\beta_j = 0$ and BF is the conditional Bayes factor for $\beta_j \neq 0$ versus $\beta_j = 0$. If BF is large (small), then $\underline{p}_j$ is small (large). Larger Bayes factors for $\beta_j \neq 0$ versus $\beta_j = 0$ lead to smaller conditional posterior probabilities that $\beta_j = 0$ and greater chances for the variable to be included in the model. If $\beta_j \neq 0$, then β_j is drawn from its conditional distribution. The main drawback to this approach is in computational speed. For larger models, SSVS and MC^3 are likely to be quicker. In this paper, we apply

[19] For example, β_j is drawn from its distribution conditional on β_p ($p = 1, 2, \ldots, k$; $p \neq j$) and σ.

Geweke's approach to select the "best" set of regressors for our mixed model. We choose this method over the others because of ease of interpretation and implementation.

3.4 Parameter Estimation in the Presence of Outliers

An observation is labeled as an outlier if it appears to be inconsistent with the rest of the data set (Barnett and Lewis, 1984). Outliers can result from recording or transmission errors or can represent actual observations which indicate the data generating mechanism is more complex than the modeler had originally thought. Outliers of the first type can be corrected, if detected, and employed in analyses. Outliers of the second type require the modeler to expand the model to explain the pattern being seen. In either case, since outliers can have a pronounced effect on the results obtained through analyses, outliers need be detected and handled appropriately.

The two major models for outliers are the slippage and mixture models. In a data set of n observations with r potential outliers, a slippage model assumes n–r observations originate from $N(\mu, \sigma^2)$ and r observations originate from a different distribution. There are two main types of slippage models, location-shift and scale-shift. In a location-shift model, the r observations come from $N(\mu + \theta, \sigma^2)$. In a scale-shift model, the r observations arise from $N(\mu, \varphi\sigma^2)$. A mixture model approach would assume the n observations are drawn from a combination of two or more distributions. Let D_1 represent the distribution that generates typical data and let D_2 be the distribution that generates the contaminated data. Beckman and Cook (1983) define a contaminant as "any observation that is not a realization from the target distribution." Then the observations originate from $(1-p)D_1 + pD_2$ where p is a fixed constant that represents the probability of being an outlier. If $p = 0$, then there are no outliers or contaminants (Iglewicz and Hoaglin, 1993).

Detection of outliers often occurs through graphical exploration of the data or during an analysis of residuals after an initial model run. Once outliers have been detected, the modeler must choose how to proceed. One extreme option would be to remove the outliers from the analysis. There are several concerns with this option. The outliers may represent an actual data phenomenon, not errors in the data. Any information the outliers would have on the issue being explored would be lost.

Other modeling techniques to incorporate outliers rely on weighted analyses of the data. Standard statistical methods, such as ordinary least squares (OLS), can be quite sensitive to outliers. Under OLS, all observations are given the same weight and the technique minimizes the sum of squared errors. In the presence of outliers, equal weights may be an incorrect approach. Weighted least squares attaches weights to each observation and minimizes the sum of weighted squared errors. Outliers receive less weight or importance in the analysis. If the variance structure of the data were known, the correct weights would be inversely proportional to the variance of each data point. However, in almost all statistical work, such knowledge of the variance structure does not exist (Everitt and Dunn, 1991).

One explanation for outliers is heteroscedasticity of the data, the variance structure of the data is not constant across observations. In a regression framework, if the data are heteroscedastic but otherwise pairwise uncorrelated, then OLS provides unbiased, but inefficient, parameter estimates. The frequentist approach to handling outliers depends on the knowledge of the variance structure. If the variance structure is known, weighted least squares can be employed to estimate the model. If the variance structure is unknown, a two-step estimation procedure, such as feasible generalized least squares (Greene, 1990), or maximum likelihood is employed. In the two-step procedure, estimates of the variance structure are obtained first, then the parameter estimates are based on these variance estimates.

To illustrate the above procedures, consider the following linear model, $\mathbf{y} = \mathbf{X}\beta + \varepsilon$, $E[\varepsilon \mid \mathbf{X}] = 0$, $Var(\varepsilon_i) = \sigma_i^2 = \sigma^2\omega_i$, where $\mathbf{y}$ is an n x 1 vector of observations, $\mathbf{X}$ is an n x 1 vector of covariates, β is an unknown scalar parameter, and ε is an n x 1 error vector. The variance component, σ^2, is unknown. The weights ω_i satisfy $\sum_{i=1}^{n}\omega_i = n$. If the disturbances are homoscedastic, then $\omega_i = 1$ for all i. Heteroscedasticity implies that the weights ω_i differ. For the regression estimates, OLS computes $\hat{\beta}_{OLS}$ as $[\mathbf{X'X}]^{-1}\mathbf{X'y}$. Let

$$\Omega = \begin{bmatrix} \omega_1 & 0 & \cdots & 0 \\ 0 & \omega_2 & & 0 \\ \vdots & & \ddots & \\ 0 & 0 & & \omega_n \end{bmatrix}.$$

If Ω is known, then the generalized least squares (GLS) estimator of β is $\hat{\beta}_{GLS} = (\mathbf{X'}\Omega^{-1}\mathbf{X})^{-1}\mathbf{X'}\Omega^{-1}\mathbf{y}$. The variances of the estimators are

$$Var(\hat{\beta}_{OLS}) = \sigma^2[\mathbf{X'X}]^{-1}[\mathbf{X'}\Omega^{-1}\mathbf{X}][\mathbf{X'X}]^{-1} = \frac{\sigma^2\sum_{i=1}^{n}X_i^4}{\left(\sum_{i=1}^{n}X_i^2\right)^2} \text{ and}$$

$$Var(\hat{\beta}_{GLS}) = \sigma^2[\mathbf{X'}\Omega^{-1}\mathbf{X}]^{-1} = \frac{\sigma^2}{n}.$$

The relative inefficiency of OLS is given by the ratio of these variances.

$$\frac{\operatorname{Var}(\hat{\beta}_{OLS})}{\operatorname{Var}(\hat{\beta}_{GLS})} = \frac{n\Sigma_{i=1}^{n} X_i^4}{\left(\Sigma_{i=1}^{n} X_i^2\right)^2} > 1.$$

If Ω is unknown, then two-step GLS or maximum likelihood could be employed to obtain parameter estimates. In two-step GLS, estimates of σ_i^2, $\hat{\sigma}_i^2$ are formed from the least squares residuals and used to compute the estimator of β,

$$\hat{\hat{\beta}} = \left[\Sigma_{i=1}^{n}\left(\frac{1}{\hat{\sigma}_i^2}\right)X_i^2\right]^{-1}\left[\Sigma_{i=1}^{n}\left(\frac{1}{\hat{\sigma}_i^2}\right)X_i y_i\right].$$

Under maximum likelihood, the estimators of β and the σ_i^2 are found by maximizing the log-likelihood function,

$$\ln L = -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\Sigma_{i=1}^{n}\left[\ln\sigma_i^2 + \frac{1}{\sigma_i^2}(y_i - \beta X_i)^2\right]$$

(Greene, 1990).

Bayesian approaches to the incorporation of outliers in analyses rely on the use of long-tailed distributions or mixture distributions to capture the information provided by the outlying points. Long-tailed distributions place substantial probability away from the center of the distribution. The family of *t*-distributions (with degrees of freedom below 10) is a classic example of long-tailed distributions. Mixture distributions are combinations of distributions weighted by probabilities. One example is a contaminated normal distribution. In a contaminated normal distribution, $\theta \sim \begin{Bmatrix} N(0,\sigma^2) & \text{with probability } \eta \\ N(0,\kappa^2\sigma^2) & \text{with probability } 1-\eta \end{Bmatrix}$, where κ^2 is a variance-inflation parameter. Box and Tiao (1968) give an early example of this modeling strategy. In this paper, we follow this mixture model structure for outliers. We choose this method over alternative approaches because of the ease of interpretation and implementation and the pattern of outliers we would expect to see from a data set such as ours.

CHAPTER 4. MODELING INVESTMENT DECISIONS THROUGH A COMPOSITE REGRESSION

4.1 The Data Set

The data employed in this analysis originate from the Individual Farm Analysis data set of the Iowa Farm Business Association. The Iowa Farm Business Association has been collecting the Individual Farm Analysis data set for a number of years. We were allowed access to the 1991-95 individual farm records. For each year the data set contains detailed production and financial information for over 1,000 Iowa farms. Income is subdivided into livestock, crop, non-cash, and non-farm sources. Expenses are categorized as cash operating, cash fixed, livestock, and non-cash expenses. Several subcategories are included in each income and expense category.

The data are collected by Iowa Farm Business Association consultants. Farm records are kept on an inventory basis under standardized accounting procedures. The value of rented land is not included in the farm's asset and liability data. Also, only the farmer's share of income and expenses are included from rented acreage. Information is also provided on the total resources on the farm, economic depreciation of assets, farm net worth, farm liabilities, and crop (or livestock) specific figures on revenues and expenses. Overall, the data set contains over 700 variables for each farm. After combining the 1991-95 data sets, we found that 667 farms had provided records for each of these years.

Of the 667 farms, 46 were missing liability and net worth information and 9 were missing machinery value information for which we could not recalculate the missing figures.

These farms were removed from the analysis. After a discussion with Mr. Duane Bennink, the supervisor of the Iowa Farm Business Association data, 22 other farms were also removed from consideration due to extremely large changes between previous end-of-year values and beginning-of-year values or to having investment ratios (the ratio of investment to the capital stock) greater than 10. This left data on 590 farms for analysis. The variable lagged age of the farm operator appears in the models used in this work. For 25 of the farms, age information was not provided by the farmer. The missing data were imputed using the age distribution for farmers in Iowa reported in the 1992 Census of Agriculture.

To provide some information on the type and size of farms in this panel data set, we present 1991-95 average values for net farm income, total farm resources (measured in dollars), total acreage, total crop acreage, total livestock sales, and total crop sales in Table 4.1. These averages were computed using 2,950 farm level observations. Over the five years, the farms in this data set have an average annual net farm income of over $45,000. On average, the farms have annual sales of nearly $220,000. The farms average 532 total

Table 4.1. 1991-95 average annual values

Variable	Average
	($)
Net Farm Income	45,789
Total Farm Resources	564,677
Total Livestock Sales	121,860
Total Crop Sales	96,832
	(acres)
Total Farm Acreage	532
Total Crop Acreage	477

acres, 477 of which is planted to a crop (or counted in federal crop programs). Crop sales account for an average of 56 percent of total sales receipts on the farm.

The farms in this data set are a self-selected sample. The farmers represented here have chosen to submit information to the Iowa Farm Business Association. To see how this self-selection might impact the analysis, we have compared the data set to figures for all Iowa farms in the 1992 Census of Agriculture. Figure 4.1 shows the comparison for farm size. The data set underrepresents both large (more than 2000 acres) and small (less than 180 acres) farms and overrepresents the farms in-between. Similar patterns emerge in the comparisons for machinery value and operator age.

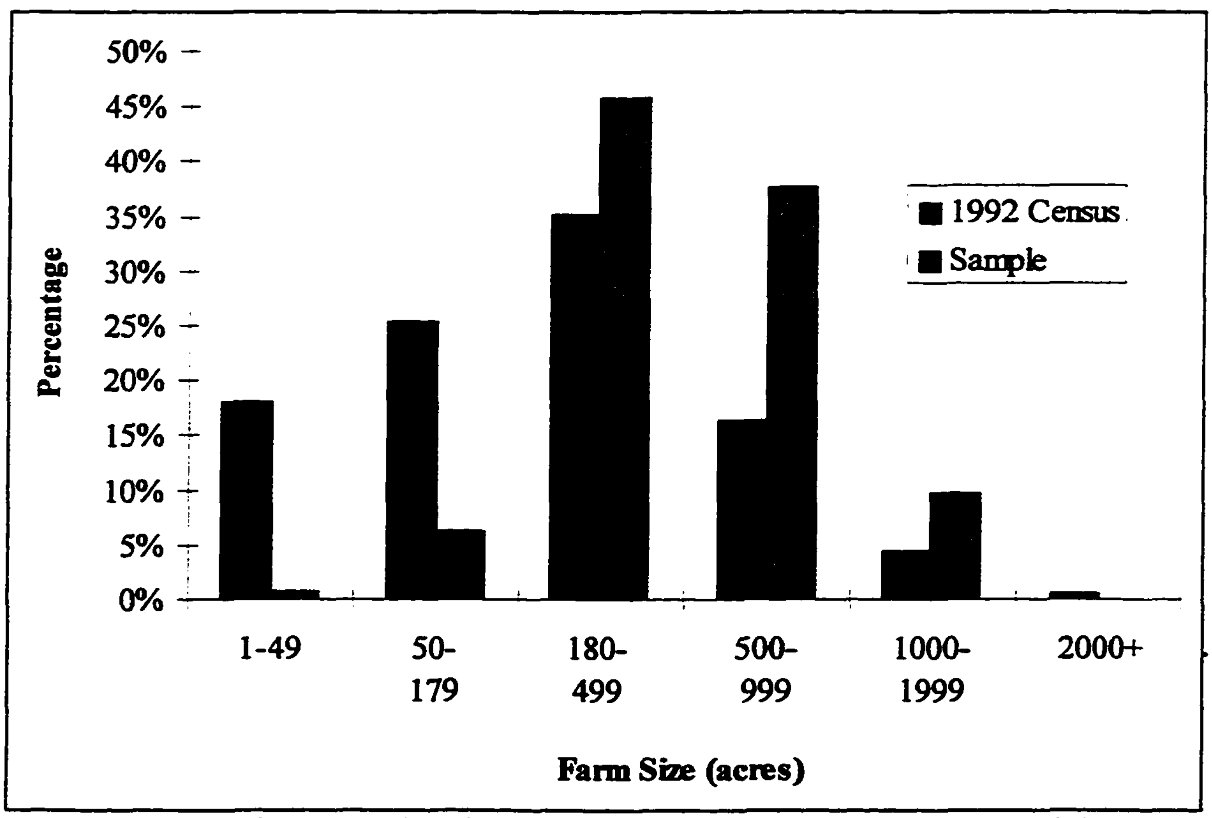


Figure 4.1. Farm size comparison between the 1992 Census of Agriculture and the sample

4.2 Exploratory Analysis

Many of the variables in the data set detail the allocation of resources to various farm activities, revenues, and expenses and thus are not relevant to this analysis. The variables chosen for inclusion in this study represent various formal investment models or serve as proxies for the 5 Cs of lending. The data set contains several financial variables for the farms. The choices of net worth and the liability figures from these financial variables is explained below.

During the process of gathering and managing the data, we began the estimation procedure by examining the data through simple graphical and statistical techniques. Such techniques can provide a quick check of the data structure and point out possible data inconsistencies.[1] The data manipulations mentioned in Section 4.1 followed an initial estimation of summary statistics on the data. The data set contains a large number of variables which could have been employed in this study. To narrow our focus, we examine the relationships among the prospective regressor candidates and the dependent variable, machinery and equipment investment. In the agricultural investment literature, several factors have been shown to affect the investment decision. The accelerator model of investment links changes in output to investment. The neoclassical model connects the user cost (or rental price) of capital to investment. Other investment studies have chosen to include a single farm financial variable, such as net worth or the debt-to-asset ratio, in their analysis and found a significant relationship between it and investment.

[1] In the first scan of the original data, we found a farmer reported to be 447 years old. Additional data indicated that the number should have been 44.

Given the many financial variables provided in our data set and the various forms these financial variables can take,[2] considerable exploratory analysis was performed. One potentially large problem with the inclusion of several farm financial variables in econometric work is multicollinearity. The accounting structure of financial variables can lead to exact multicollinearity. For example, net worth is equal to the difference between total assets and total liabilities. After examining correlations among several financial variables both in level and ratio formats, it was found that using the variables in the standard financial ratios would increase the likelihood of collinearity problems. Thus, we proceed with the financial variables in level form. To maintain consistency across the two approaches employed in this study, several cash flow measures were removed from consideration as regressors.[3]

The original structure of the model included individual farm intercepts. After some initial examination of the data, we decided to explore whether investment may have an autoregressive component. We estimated both a simple regression and a random effects model[4] of investment with lagged investment and squared lagged investment. Both models suggested the inclusion of an autoregressive component in the investment model.

In Table 4.2, the summary statistics for the study variables are given. The data set contains 1770 observations. All monetary values are deflated. Output is measured by total

---

[2] For example, current liabilities could be brought into a model in level form or through a ratio form, such as the current ratio (current assets/current liabilities) or the current debt ratio (current liabilities/total liabilities).

[3] In the Euler equation approach to investment modeling, it is the farmer's cash flow that is modeled. Thus, when a cash flow variable is added to the structure used in this model, twice-lagged variables are required to proceed. Given the short time frame of the data set, we decided against employing such variables in the analysis.

[4] The random effects are year intercepts to capture aggregate economic events that impact all farmers.

cash farm income. The cost of capital[5] is an index representing the price at which capital may be obtained. The farms within the data set vary a great deal in size. In general, the farms were growing over the time period, as can be seen from the average investment and output change figures. But some farms did go through quite dramatic business contractions. Operators ages ranged from the lower 20s to nearly 80.

Table 4.2. Summary statistics

Variable	Mean	Standard Deviation	Minimum	Maximum
Investment ($I^N_{i,t}$)	5001.21	23193.67	-117937.22	149683.44
Change in output ($\Delta Q_{i,t}$)	10415.77	69654.32	-915386.52	480314.47
Value of short-term assets ($V_{i,t}$)	170280.24	132051.93	0.00	1514305.94
Cost of capital ($C_{i,t}$)	22.65	1.85	19.76	26.73
Lagged operator age ($AGE_{i,t-1}$)	46.90	10.97	23.00	79.00
Lagged total liabilities ($TL_{i,t-1}$)	170890.85	187674.57	0.00	1211338.00
Lagged net worth ($NW_{i,t-1}$)	450020.65	378784.87	-100880.00	2360768.00
Lagged current liabilities ($CL_{i,t-1}$)	62650.95	88206.25	0.00	799883.00
Lagged machinery value ($K_{i,t-1}$)	99579.27	69055.56	1205.10	444220.00
Lagged investment ($I^N_{i,t-1}$)	7103.65	22656.15	-128708.96	149683.44

Disinvestment (negative investment) was reported 51.2 percent of the time, with extreme disinvestment (greater than \$50,000) occurring 0.7 percent of the time. Annual investments of over \$100,000 took place in 0.8 percent of the observations. In 81.7 percent of the observations, farmers reported that they faced some level of debt and at least part of the debt was due within the next year 72.4 percent of the time. Four farmers reported liabilities above \$1 million. Nearly ten percent of the observations showed farms with a net worth exceeding \$1 million, while 1.2 percent displayed a negative net worth. Only one

[5] The cost of capital is explained in more detail in the next section (Section 4.3).

percent of the observations reported farm machinery values below \$10,000. Meanwhile, 1.7 percent of the observations showed farm machinery values above \$300,000.

As discussed earlier, the presence of multicollinearity among the variables is a distinct possibility when examining financial data. Table 4.3 displays the correlation matrix among the variables in this analysis. As might be expected, the highest degree of correlation is between the level of current and total liabilities on the farm. The value of short-term assets, net worth, and the value of machinery on the farm are also quite correlated among themselves and with the liability figures.

Table 4.3. Correlation matrix

	$I^N_{i,t}$	$\Delta Q_{i,t}$	$V_{i,t}$	$C_{i,t}$	$AGE_{i,t-1}$	$TL_{i,t-1}$	$NW_{i,t-1}$	$CL_{i,t-1}$	$K_{i,t-1}$
$\Delta Q_{i,t}$	0.07								
$V_{i,t}$	0.14	0.07							
$C_{i,t}$	-0.04	0.31	0.23						
$AGE_{i,t-1}$	-0.10	-0.06	0.06	0.02					
$TL_{i,t-1}$	0.05	0.08	0.45	0.04	-0.03				
$NW_{i,t-1}$	0.05	0.00	0.60	0.21	0.35	0.13			
$CL_{i,t-1}$	0.06	0.09	0.53	0.05	-0.03	0.77	0.10		
$K_{i,t-1}$	-0.01	0.06	0.52	0.18	0.03	0.47	0.52	0.39	
$I^N_{i,t-1}$	0.03	0.04	0.12	0.05	-0.07	0.09	0.10	0.09	0.33

To show the distribution of investment values, we have formed a histogram of the values in Figure 4.2. As the figure shows, most of the investment moves made by these farmers are made to maintain the capital stock. Nearly 80 percent of the investment totals are between -\$20,000 and \$20,000. The average investment rate is five percent of the capital stock. The distribution of investment values is somewhat skewed. The mean value is just over \$5,000, while the median value is -\$287.

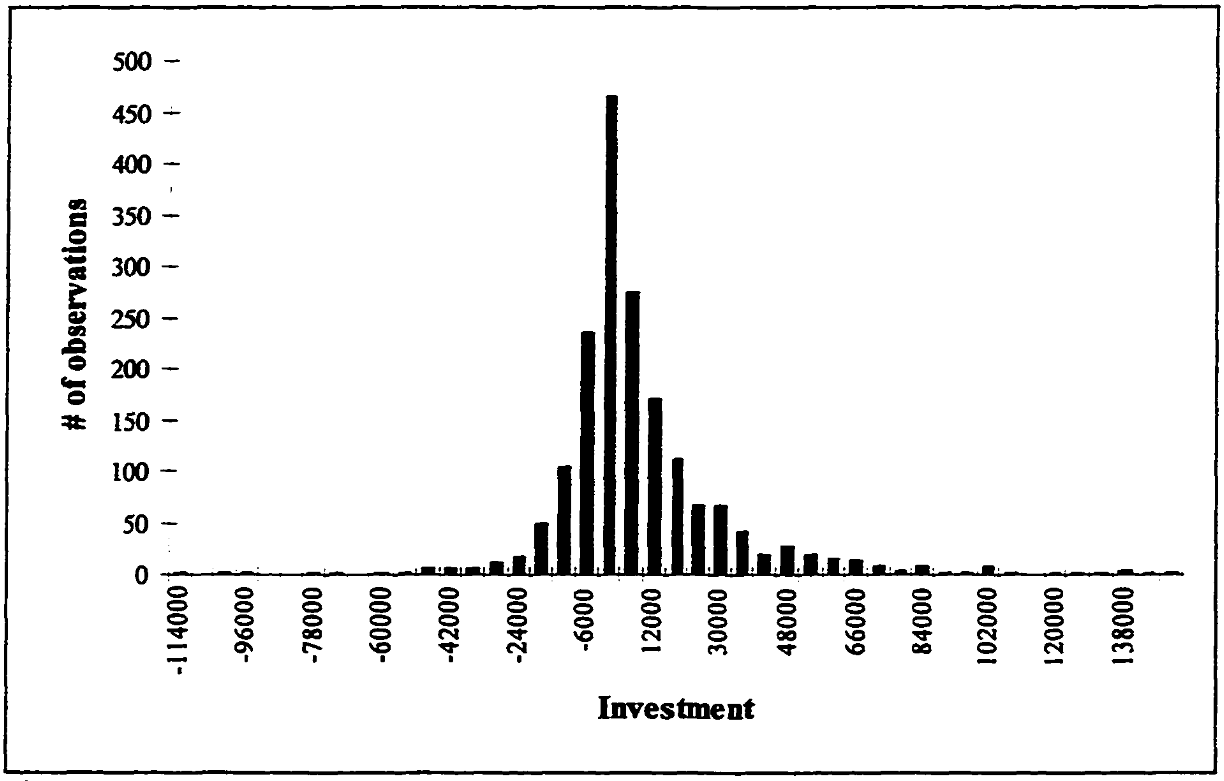


Figure 4.2. Histogram of investment values

The histogram of investment values also shows that outliers may exist in the data set. There are several large positive and negative investment values. In a typical regression analysis, these points could have a significant impact on the results. To explore these points further, we have produced scatter plots of the regressors versus investment (Figures 4.3-4.11). In each of these plots, the regressor is shown on the horizontal axis and investment is shown on the vertical axis. After reviewing these graphs, we decided to include an outlier detection component in the final model. Since the graphs indicate the possibility of outliers with extremely high and low investment values, we have chosen a variance-inflation model to capture any outliers.

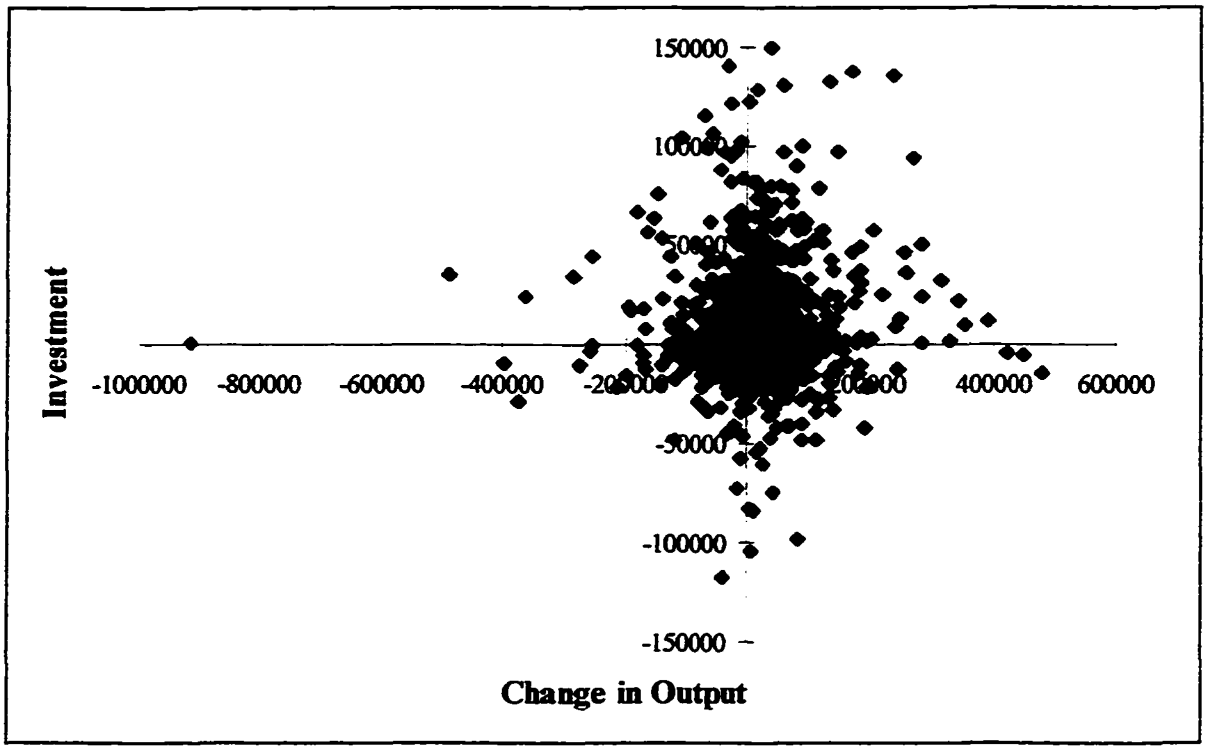


Figure 4.3. Scatter plot of the change in output vs. investment

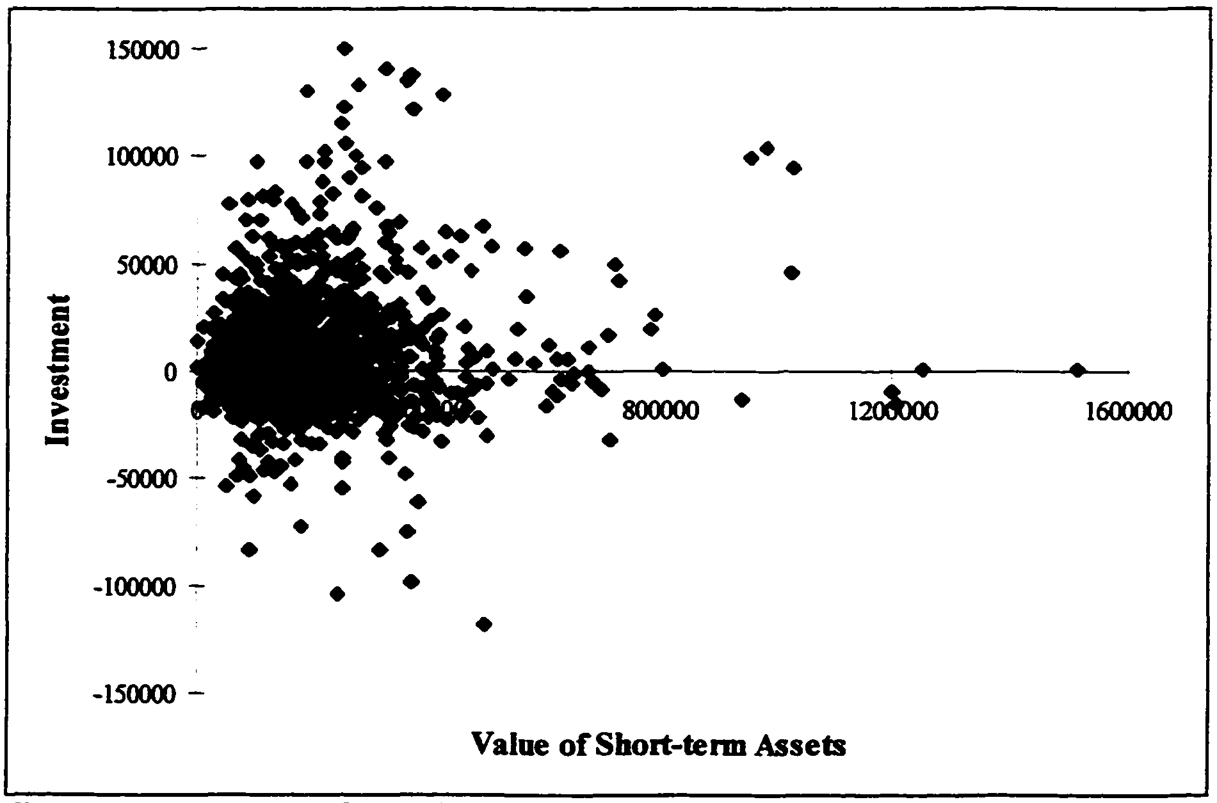


Figure 4.4. Scatter plot of the value of short-term assets vs. investment

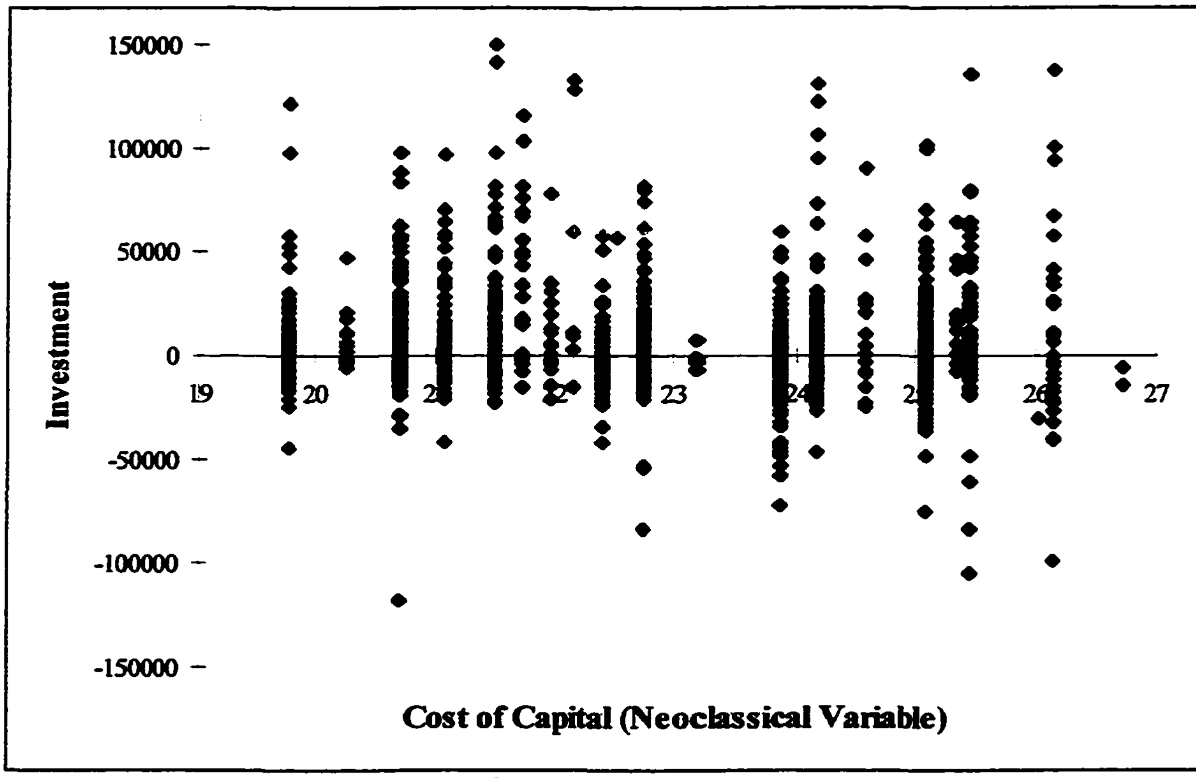


Figure 4.5. Scatter plot of the cost of capital vs. investment

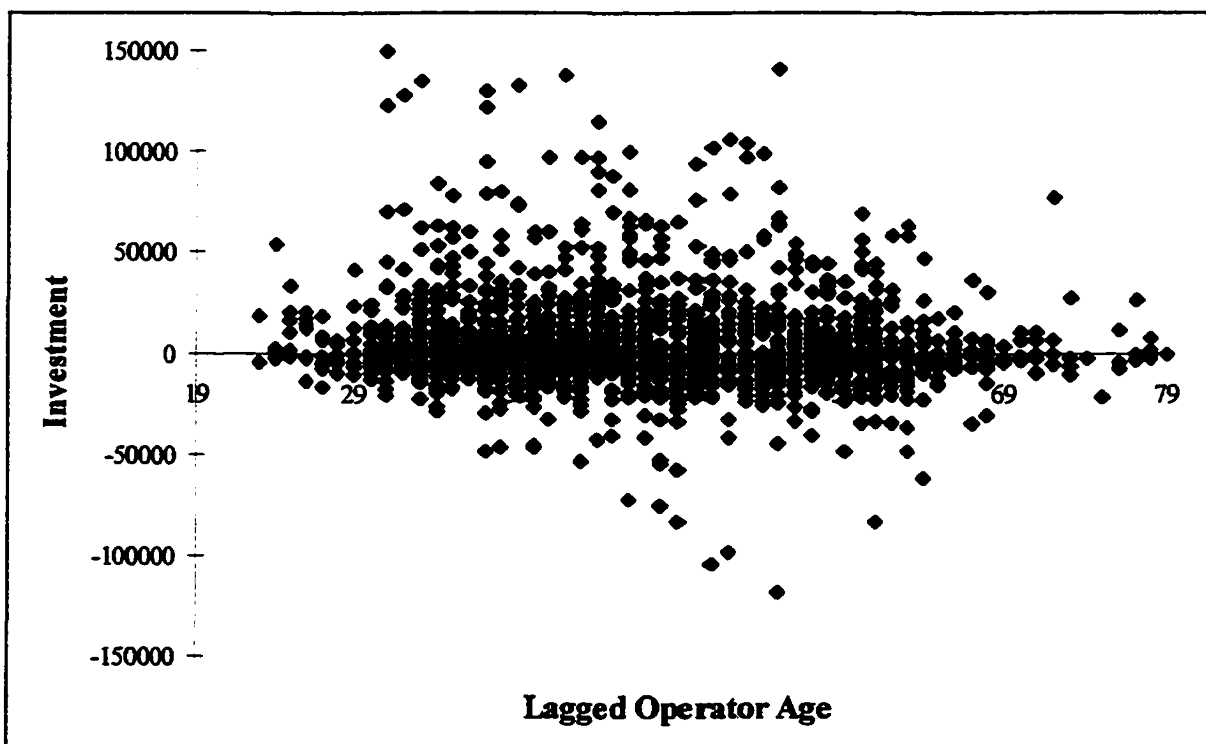


Figure 4.6. Scatter plot of lagged operator age vs. investment

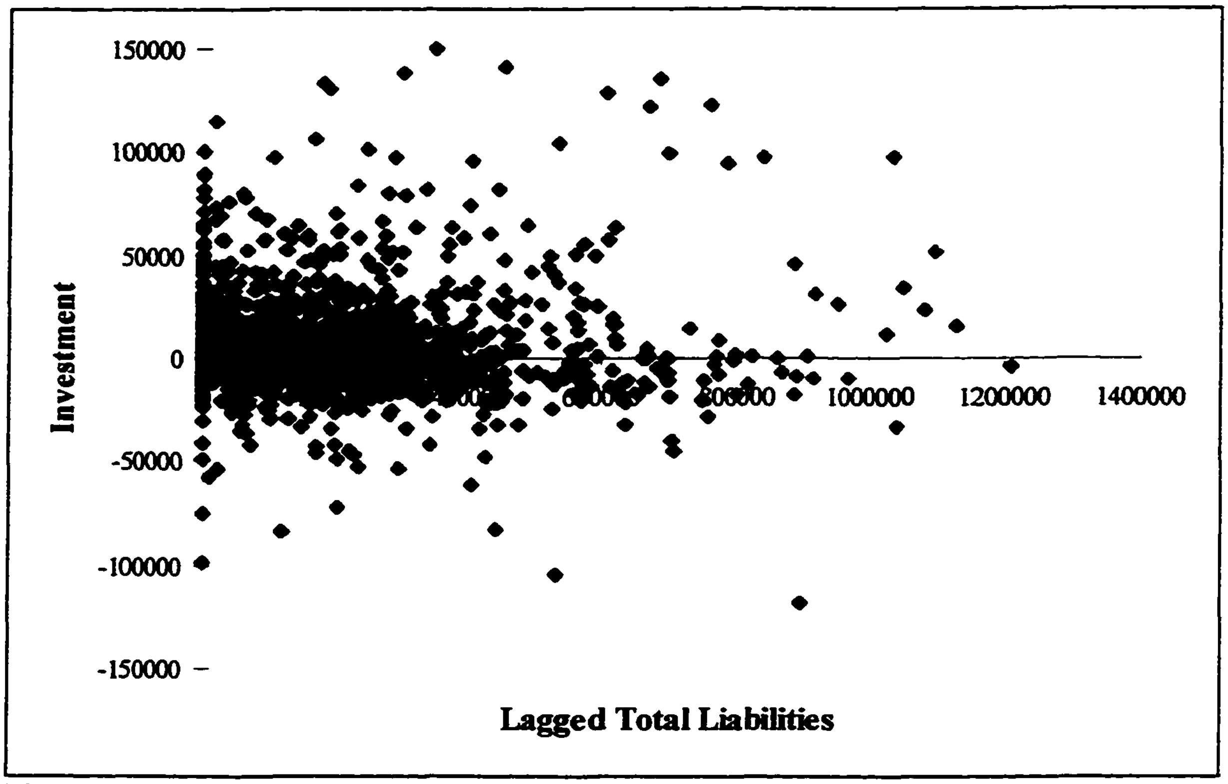


Figure 4.7. Scatter plot of lagged total liabilities vs. investment

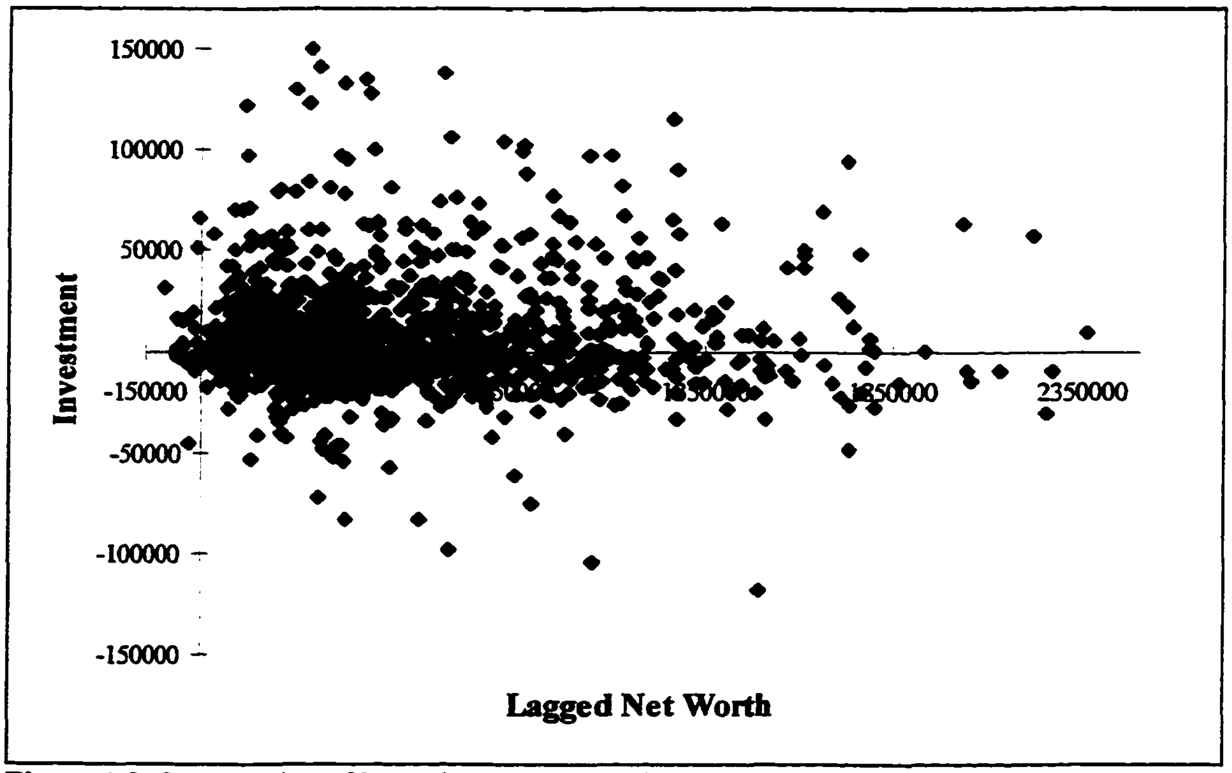


Figure 4.8. Scatter plot of lagged net worth vs. investment

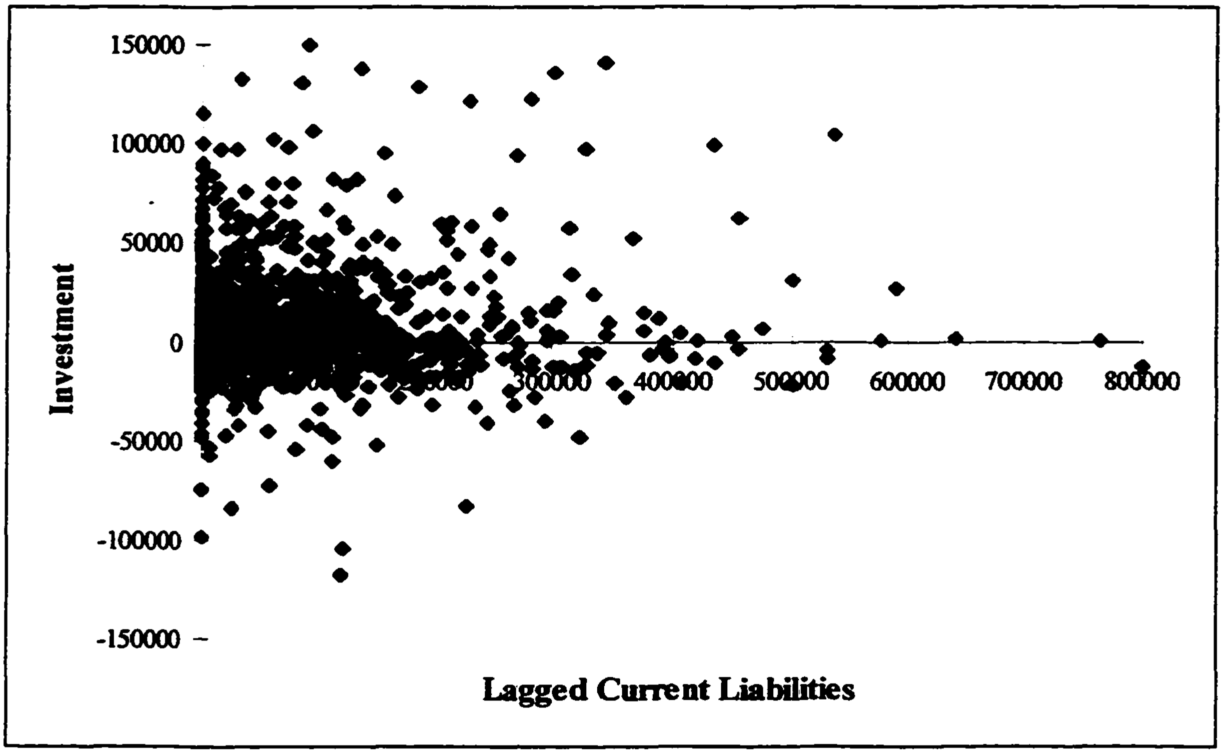


Figure 4.9. Scatter plot of lagged current liabilities vs. investment

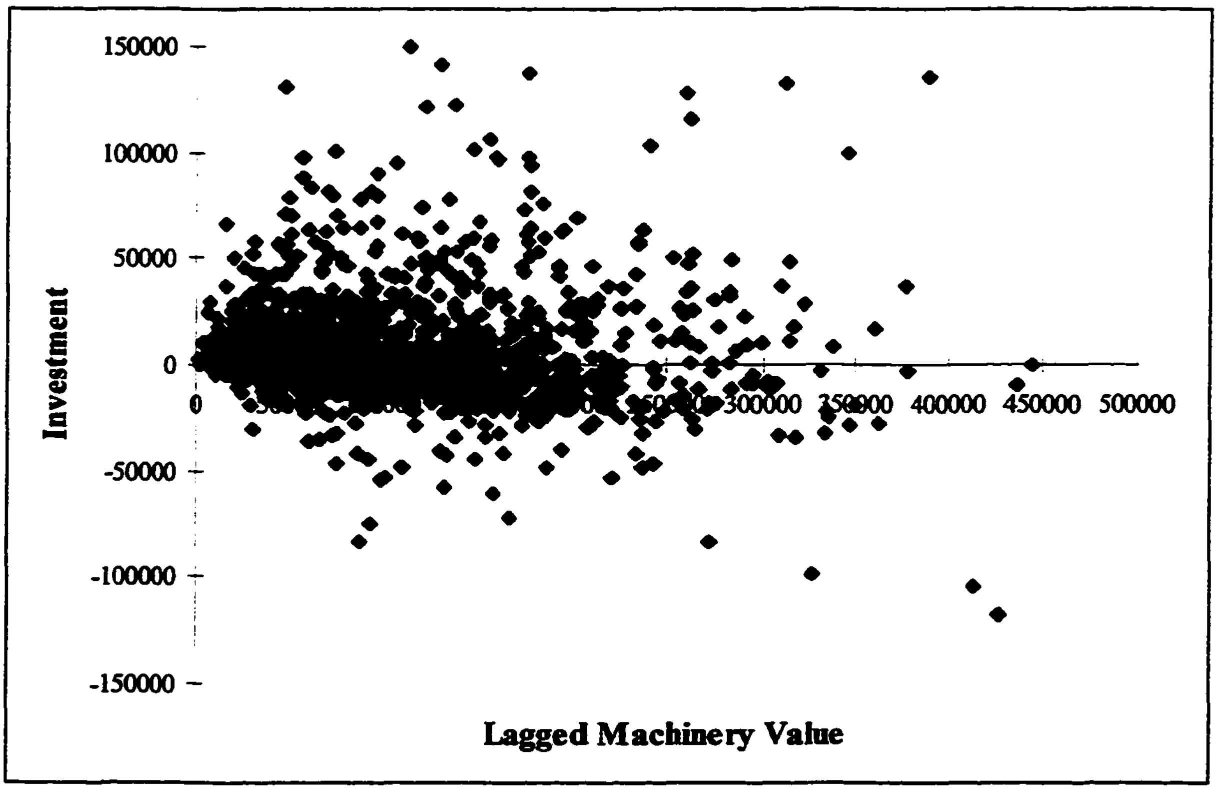


Figure 4.10. Scatter plot of lagged machinery value vs. investment

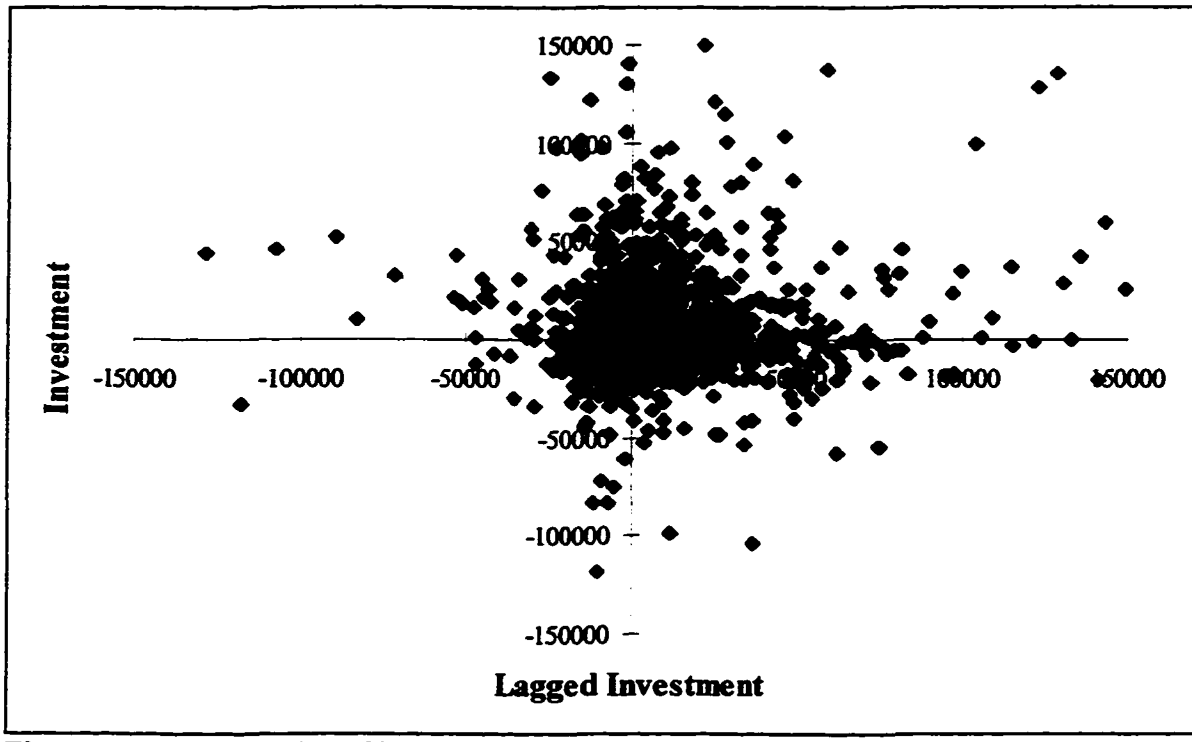


Figure 4.11. Scatter plot of lagged vs. current investment

4.3 The Model

In this analysis, we examine agricultural investment in machinery and equipment for Iowa farms. We form an investment regression model by combining aspects from the accelerator and neoclassical investment models with other possibly influential variables such as internal financial variables. Inclusion of the internal finance variables can be justified through claims that farms face financial constraints in their investment decisions. Financial constraints would arise if there is an asymmetric information problem between the prospective lender and the farmer or if there are substantial transactions costs to obtaining outside financing.

Let subscripts i and t represent farm and year, respectively. Net investment ($I^N_{i,t}$) is measured as the difference between the value of machinery and equipment at the beginning and the end of the year. The investment regression includes variables representing changes in output ($\Delta Q_{i,t}$), the value of short-term assets ($V_{i,t}$), the cost of capital ($C_{i,t}$), owner net worth ($NW_{i,t-1}$), total farm liabilities ($TL_{i,t-1}$), current liabilities ($CL_{i,t-1}$), and operator age ($AGE_{i,t-1}$), along with the previous values of machinery and equipment ($K_{i,t-1}$) and investment ($I^N_{i,t-1}$). The value of short-term assets, net worth, farm liabilities, current liabilities, and operator age are chosen to represent the 5 Cs of the farmer credit situation. The form of the regression is

$$I^N_{i,t} = \beta_0 + \Sigma\beta_k X_{i,k,t} + \Sigma\beta_j (X_{i,k,t} X_{i,p,t}) + \varepsilon_{i,t}$$

where the β_h are unknown parameters to be estimated, the Xs represent the various regressors, and $\varepsilon_{i,t}$ is the regression error. The model has nine main effects and 45 cross effects. The user cost of capital ($C_{i,t}$) originates from the neoclassical model of investment. The equation for computing $C_{i,t}$ is given by

$$C_{i,t} = \left(\frac{p^K_t}{1-m_{i,t}}\right)\left[(1-m_{i,t})\delta + r_t - \left(\frac{p^K_t - p^K_{t-1}}{p^K_t}\right)\right],$$

where p^K represents the price of new capital, $m_{i,t}$ is the marginal tax rate, δ is the capital depreciation rate, and r_t is represents the interest rate.

We employ both classical and Bayesian techniques to estimate the parameters in the regression model. Due to the panel structure of the data, classical methods would include the addition of either random or fixed effects to the regression. In this case, a typical one-way fixed effect model would append separate intercepts for each year to the equation. The resulting regression model is given by

$$I^{N}_{i,t} = \beta_0 + \Sigma\beta_k X_{i,k,t} + \Sigma\beta_j (X_{i,k,t} X_{i,p,t}) + y_t + \varepsilon_{i,t}\,,$$

where the y_t are year intercepts. A random effects model for the regression includes year intercepts and assumes that the y_t have the following properties: $E[y_t] = 0$, $Var(y_t) = \sigma_y^2$, and $Cov(y_t, \varepsilon_{i,t}) = 0$.

For the Bayesian approach, we formulate our regression model as a hierarchical normal linear model. We use Geweke's (1996) variable selection method to choose regressors. Because exploratory analyses of the data indicate that outliers may be present, we model the residuals $\varepsilon_{i,t}$ as coming from a contaminated normal distribution, as described in Section 3.4. For the investment model given above, we assume that $y_t \sim$ iid $N(0, \sigma_y^2)$ for all t and $\varepsilon_{i,t} \sim \begin{Bmatrix} N(0,\sigma_\varepsilon^2) & \text{with probability } \eta \\ N(0,\kappa^2\sigma_\varepsilon^2) & \text{with probability } 1-\eta \end{Bmatrix}$ where κ^2 is a variance-inflation parameter. Let $\boldsymbol{\beta}' = (\beta_0, \beta_1, \ldots, \beta_{54})$, $\mathbf{X}_{i,t}$ be the row vector of regressors for the i,t[th] observation, and $\theta_{i,t}$ be one if $\varepsilon_{i,t} \sim N(0, \sigma_\varepsilon^2)$ and zero otherwise. The distributional assumptions, combined with the regression equation, imply that $I^{N}_{i,t} \mid y_t, \boldsymbol{\beta}, \mathbf{X}_{i,t}, \sigma_\varepsilon^2, \sigma_y^2, \theta_{i,t} \sim N(y_t + \mathbf{X}_{i,t}\boldsymbol{\beta}, \sigma_\varepsilon^2(\theta_{i,t} + \kappa^2 - \kappa^2\theta_{i,t}))$.

4.4 Prior Distributions

In the Bayesian framework, y_t, $\boldsymbol{\beta}$, and σ_ε^2 are the parameters of the model, and η, $\theta_{i,t}$, and σ_y^2 are the hyperparameters. The joint posterior distribution of all parameters in the model is obtained by combining the likelihood function with prior distributions for the parameters and hyperparameters. For σ_ε^2, we chose a noninformative prior distribution, $p(\sigma_\varepsilon^2) \propto 1/\sigma_\varepsilon^2$. For σ_y^2, we chose an informative prior distribution, $\sigma_y^2 \sim \text{Inverse-}\chi^2(n_0, \sigma_0^2)$. An informative prior for σ_y^2 was chosen because the data contain little information about the effect of time on investment. The effect of the informative prior can be thought of as adding n_0 observations with an average squared deviation of σ_0^2 to the analysis of σ_y^2. A priori, σ_ε^2 and σ_y^2 were modeled as independent parameters. Thus, $p(\sigma_\varepsilon^2, \sigma_y^2) = p(\sigma_\varepsilon^2)p(\sigma_y^2)$.

The prior distributions for η and $\theta_{i,t}$ are Beta distributions,

$$p(\eta \mid \gamma, \varphi) \propto \eta^{(\gamma-1)}(1-\eta)^{(\varphi-1)} \text{ and } p(\theta_{i,t} \mid \eta) \propto \eta^{\theta_{i,t}}(1-\eta)^{(1-\theta_{i,t})},$$

respectively. We set the hyperparameters γ and φ to values that reflect our prior beliefs on the proportion of potential outliers that may be present in the data set. To assess the sensitivity of the results to the priors for the variable selection and outlier detection components of the model, we estimate the model under ten various combinations of priors. For the outlier detection component, we use three sets of prior distributions for the outlier detection parameters: a prior strongly suggesting that 10 percent of the observations are outliers ($\gamma = 18$, $\varphi = 2$), a weaker flat prior suggesting that 50 percent of the observations are outliers ($\gamma = 1$, $\varphi = 1$), and a prior strongly suggesting that 90 percent of the observations are outliers ($\gamma = 2$, $\varphi = 18$). For the variable selection component, we also employ three priors: a prior suggesting a 10 percent probability that each regressor (main and cross effects) is included in the model ($\underline{p}_j = 0.9$); a prior suggesting a 50 percent probability ($\underline{p}_j = 0.5$); and a prior suggesting a 90 percent probability ($\underline{p}_j = 0.1$). The nine combinations of these priors are all examined, along with a run assuming that the data set contains no outliers and that each variable has a 50 percent prior probability of being included in the model. In total, 10 estimations are performed.

4.5 The Gibbs Sampler for the Mixed Model with Variable Selection and Outlier Detection

Given the model specified in Section 4.3 and the prior distributions for the parameters specified in Section 4.4, the Gibbs sampler for this problem has six major components:

1) simulation of the main outlier distribution parameter, η,
2) simulation of the main error variance, σ_ε^2,
3) simulation of the annual random effect variance, σ_y^2,
4) simulation of the annual random effects, y_t,
5) simulation of the individual observation outlier detection parameter, $\theta_{i,t}$, and
6) simulation of the parameter vector, $\boldsymbol{\beta}$.

We have designed our Gibbs sampler to handle these simulations in the order given above. The conditional distribution $p(\eta \mid I^N_{i,t}, y_t, \boldsymbol{\beta}, \mathbf{X}_{i,t}, \sigma_\varepsilon^2, \sigma_y^2, \theta_{i,t})$ is a Beta distribution,

$$\eta \mid I^N_{i,t}, y_t, \boldsymbol{\beta}, \mathbf{X}_{i,t}, \sigma_\varepsilon^2, \sigma_y^2, \theta_{i,t} \sim \text{Beta}\left(\textstyle\sum_{i=1}^{n}\sum_{t=1}^{T}\theta_{i,t} + \gamma,\ nT - \sum_{i=1}^{n}\sum_{t=1}^{T}\theta_{i,t} + \varphi\right),$$

where n is the number of farms, T is the number of years, and the hyperparameters γ and φ are values that reflect our prior beliefs[6] on the proportion of outliers that may be present in the data set.

The conditional distribution for the main error variance, σ_ε^2, is given by an Inverse-Gamma distribution,

---

[6] We examine three sets of priors, as explained in the previous section.

$$\sigma_\varepsilon^2 \mid I^N_{i,t}, y_t, \beta, \mathbf{X}_{i,t}, \eta, \sigma_y^2, \theta_{i,t} \sim \text{Inverse-Gamma}\left(0.5nT,\ 0.5\Sigma_{i=1}^{n}\Sigma_{t=1}^{T}\frac{\left(I^N_{i,t}-\mathbf{X}_{i,t}\beta-y_t\right)^2}{\left(\theta_{i,t}+\kappa^2-\kappa^2\theta_{i,t}\right)}\right),$$

where κ^2 is the variance-inflation parameter. The conditional distribution for the annual random effects variance follows a similar structure,

$$\sigma_y^2 \mid I^N_{i,t}, y_t, \beta, \mathbf{X}_{i,t}, \eta, \sigma_\varepsilon^2, \theta_{i,t} \sim \text{Inverse-Gamma}\left(0.5(T+n_0),\ 0.5\left(n_0\sigma_0^2+\Sigma_{t=1}^{T}y_t^2\right)\right),$$

where n_0 and σ_0^2 are hyperparameters from the informative prior placed on σ_y^2.

For the annual random effects, each effect has a Normal conditional distribution,

$$y_t \mid I^N_{i,t}, \sigma_y^2, \beta, \mathbf{X}_{i,t}, \eta, \sigma_\varepsilon^2, \theta_{i,t} \sim \left(\left(\frac{\sigma_y^2}{\sigma_\varepsilon^2+\sigma_y^2W}\right)\Sigma_{i=1}^{n}\frac{\left(I^N_{i,t}-\mathbf{X}_{i,t}\beta\right)}{\left(\theta_{i,t}+\kappa^2-\kappa^2\theta_{i,t}\right)},\left(\frac{1}{\sigma_\varepsilon^2}W+\frac{1}{\sigma_y^2}\right)^{-1}\right)$$

where $W=\Sigma_{i=1}^{n}\frac{1}{\left(\theta_{i,t}+\kappa^2-\kappa^2\theta_{i,t}\right)}$. The conditional distributions for the outlier indicators are given by $p(\theta_{i,t} \mid I^N_{i,t}, y_t, \beta, \mathbf{X}_{i,t}, \eta, \sigma_\varepsilon^2, \sigma_y^2) \propto \left(\frac{\kappa\eta}{1-\eta}\right)^{\theta_{i,t}}\exp\left(\frac{-1}{2\sigma_\varepsilon^2}\frac{\left(I^N_{i,t}-\mathbf{X}_{i,t}\beta-y_t\right)^2}{\left(\theta_{i,t}+\kappa^2-\kappa^2\theta_{i,t}\right)}\right)$.

Since $\theta_{i,t}$ can only take on the values of zero and one, we can see that

$$p(\theta_{i,t}=0 \mid I^N_{i,t}, y_t, \beta, \mathbf{X}_{i,t}, \eta, \sigma_\varepsilon^2, \sigma_y^2) = \frac{p_0}{p_0+p_1} \text{ and}$$

$$p(\theta_{i,t}=1 \mid I^N_{i,t}, y_t, \beta, \mathbf{X}_{i,t}, \eta, \sigma_\varepsilon^2, \sigma_y^2) = \frac{p_1}{p_0+p_1}$$

where $p_0=\exp\left(\frac{-\left(I^N_{i,t}-\mathbf{X}_{i,t}\beta-y_t\right)^2}{2\kappa^2\sigma_\varepsilon^2}\right)$ and $p_1=\left(\frac{\kappa\eta}{1-\eta}\right)\exp\left(\frac{-1}{2\sigma_\varepsilon^2}\left(I^N_{i,t}-\mathbf{X}_{i,t}\beta-y_t\right)^2\right)$.

Again, for the regression parameters, we are following the procedure outlined by Geweke (1996). Given $\theta_{i,t}$, $I^N_{i,t}$, y_t, $X_{i,t}$, η, σ_ε^2, σ_y^2, and β_j ($j \neq k$), the conditional distribution of β_k originates from the simplified model:

$$z_{i,t} = \beta_k X_{i,t,k} + \varepsilon_{i,t}, \text{ where } \varepsilon_{i,t} \sim \begin{Bmatrix} N(0,\sigma_\varepsilon^2) & \text{with probability } \eta \\ N(0,\kappa^2\sigma_\varepsilon^2) & \text{with probability } 1-\eta \end{Bmatrix}$$

where $z_{i,t} = I^N_{i,t} - \sum_{j\neq k}\beta_j X_{i,t,j} - y_t$. Assuming a Normal prior on β_k, $p(\beta_k) \propto \exp\left(\frac{-\beta_k^2}{2\tau_k^2}\right)$.

Then, given the prior probability that $\beta_k = 0$, $\underline{p}_k$, the posterior probability that $\beta_k = 0$ is given by

$$\bar{p}_k = \frac{\underline{p}_k}{\left(\underline{p}_k + \left[1-\underline{p}_k\right]BF\right)},$$

where BF, the conditional Bayes factor in favor of $\beta_k \neq 0$ versus $\beta_k = 0$, is given by

$$BF = \left(\frac{1}{\tau_k}\left(\frac{1}{\sigma_\varepsilon^2}\sum_{i=1}^n\sum_{t=1}^T\frac{X_{i,t,k}^2}{\left(\theta_{i,t}+\kappa^2-\kappa^2\theta_{i,t}\right)}-\frac{1}{\tau_k^2}\right)^{-0.5}\right)$$

$$\exp\left(0.5\left(\frac{1}{\sigma_\varepsilon^2}\sum_{i=1}^n\sum_{t=1}^T\frac{X_{i,t,k}z_{i,t}}{\left(\theta_{i,t}+\kappa^2-\kappa^2\theta_{i,t}\right)}\right)^2\left(\frac{1}{\sigma_\varepsilon^2}\sum_{i=1}^n\sum_{t=1}^T\frac{X_{i,t,k}^2}{\left(\theta_{i,t}+\kappa^2-\kappa^2\theta_{i,t}\right)}-\frac{1}{\tau_k^2}\right)^{-1}\right).$$

The posterior probability $\beta_k \neq 0$ is equal to $1 - \bar{p}_k$.

4.6 Computational Strategy

As stated before, the analysis consists of ten separate Bayesian estimation runs. Within each run, the Gibbs sampler will simulate four chains of 12,000 iterations each.

Thus, each estimation will contain 48,000 draws. The variance-inflation parameter, κ, is set at four for all runs involving outlier detection. For the prior on σ_y^2, we chose $\sigma_y^2 \sim$ Inverse-$\chi^2(20, 0.45)$, effectively adding 20 observations with an average squared deviation of 0.45 to the analysis of σ_y^2. An intercept is always included in the model and the prior standard deviation (τ) for it is set at 0.7. The prior standard deviations for the other regression parameters are set at one. The other prior hyperparameter settings are given in Section 4.4

Starting values are chosen systematically just for convenience. Doing this does not present a problem as long as the chains are "long enough" so that the runs converge. The properties of the Gibbs sampler[7] imply that the chains will have a unique stationary distribution and that distribution will be the target distribution. The first half of each chain (6,000 iterations) is discarded as a burn-in procedure. Convergence is monitored by Gelman and Rubin's R-statistic, $\sqrt{\hat{R}}$.[8] In checking convergence, we examined all parameters including the random effects, the variance components, and the outlier distribution parameter.

The simulation programs are written in C++ and are complied by Borland C++ Builder 3. The distribution subroutines are C++ programs contained in the SUM module of the M++ Version 7.0 libraries from Dyad Software Corporation. A typical run would last seven hours on a Micron personal computer with a Pentium 166 MHz chip and 48 megabytes of RAM.

[7] These properties are outlined in Section 3.2.
[8] This statistic is described in Section 3.2.

CHAPTER 5. COMPOSITE REGRESSION RESULTS

5.1 The Classical Mixed Model Results

We first estimate the parameters in the random effects model using classical methods. The random effects are represented by annual intercepts. The parameter estimates are given in Table 5.1. The model was fitted on a personal computer using SAS 6.12 for Windows. Several of the main effects are significant. Changes in output, the value of short-term assets, and net worth have a direct relationship with investment, while operator age and the value of machinery and equipment have an inverse relationship with investment. For the squared terms, only the cost of capital is significant. Of the 36 cross effects, seven have parameter estimates significantly different from zero. The variance estimates indicate that the residual error is nearly seventeen times more variable than the random effects.

From these estimates we have calculated the expected change in investment related to a one unit change in each regressor and the elasticity $\left(\frac{\partial y}{\partial x}\frac{x}{y}\right)$ for each regressor to examine the absolute and relative impacts. Given the skewness in the investment data, we calculate these measures at both the mean and median values for all variables. Table 5.2 shows the mean and median values of investment and the regressor variables. In almost every case, the mean value is larger than the median value. The median value of current investment is negative, indicating real disinvestment on the farm. The change in the sign of the investment variable from the mean value to the median value will cause the elasticities and expected changes to also change signs. Table 5.3 displays the expected changes and

Table 5.1. Mixed model results

Effect	Estimate	Standard Error	t	Pr > |t|
Intercept	0.73	0.33	2.18	0.16
$\Delta Q_{i,t}$	0.039	0.011	3.50	0.00
$V_{i,t}$	0.0377	0.0098	3.86	0.00
$C_{i,t}$	0.058	0.064	0.90	0.37
$AGE_{i,t-1}$	-0.0342	0.0062	-5.51	0.00
$TL_{i,t-1}$	0.0017	0.0070	0.24	0.81
$NW_{i,t-1}$	0.0070	0.0032	2.19	0.03
$CL_{i,t-1}$	0.007	0.017	0.41	0.68
$K_{i,t-1}$	-0.050	0.015	-3.30	0.00
$I^N_{i,t-1}$	-0.051	0.032	-1.61	0.11
$\Delta Q_{i,t}^2$	0.00015	0.00041	0.36	0.72
$V_{i,t}^2$	0.00019	0.00036	0.52	0.60
$C_{i,t}^2$	-0.056	0.024	-2.32	0.02
$AGE_{i,t-1}^2$	0.00011	0.00044	0.25	0.80
$TL_{i,t-1}^2$	-0.00003	0.00024	-0.14	0.89
$NW_{i,t-1}^2$	-0.000021	0.000051	-0.41	0.69
$CL_{i,t-1}^2$	-0.0017	0.0010	-1.67	0.10
$K_{i,t-1}^2$	-0.0024	0.0015	-1.66	0.10
${I^N_{i,t-1}}^2$	0.0081	0.0049	1.64	0.10
$\Delta Q_{i,t}*V_{i,t}$	-0.00145	0.00061	-2.38	0.02
$\Delta Q_{i,t}*C_{i,t}$	0.0005	0.0050	0.10	0.92
$\Delta Q_{i,t}*AGE_{i,t-1}$	-0.00175	0.00092	-1.91	0.06
$\Delta Q_{i,t}*TL_{i,t-1}$	0.00007	0.00058	0.13	0.90
$\Delta Q_{i,t}*NW_{i,t-1}$	0.00052	0.00028	1.86	0.06
$\Delta Q_{i,t}*CL_{i,t-1}$	0.00005	0.00093	0.05	0.96
$\Delta Q_{i,t}*K_{i,t-1}$	-0.0002	0.0017	-0.13	0.90
$\Delta Q_{i,t}*I^N_{i,t-1}$	0.0035	0.0031	1.10	0.27
$V_{i,t}*C_{i,t}$	0.0059	0.0035	1.66	0.10
$V_{i,t}*AGE_{i,t-1}$	0.00089	0.00073	1.21	0.23
$V_{i,t}*TL_{i,t-1}$	0.00084	0.00048	1.72	0.08
$V_{i,t}*NW_{i,t-1}$	-0.00047	0.00023	-2.06	0.04
$V_{i,t}*CL_{i,t-1}$	-0.00063	0.00092	-0.68	0.49
$V_{i,t}*K_{i,t-1}$	0.0000	0.0011	0.01	0.99
$V_{i,t}*I^N_{i,t-1}$	0.0015	0.0027	0.57	0.57
$C_{i,t}*AGE_{i,t-1}$	-0.0016	0.0032	-0.49	0.62
$C_{i,t}*TL_{i,t-1}$	-0.0011	0.0027	-0.41	0.68
$C_{i,t}*NW_{i,t-1}$	0.0006	0.0013	0.50	0.62
$C_{i,t}*CL_{i,t-1}$	-0.0038	0.0059	-0.64	0.52

Table 5.1. (continued)

Effect	Estimate	Standard Error	t	Pr > |t|
$C_{i,t}*K_{i,t-1}$	-0.0198	0.0064	-3.09	0.00
$C_{i,t}*I^N_{i,t-1}$	0.009	0.016	0.58	0.57
$AGE_{i,t-1}*TL_{i,t-1}$	-0.00071	0.00059	-1.22	0.22
$AGE_{i,t-1}*NW_{i,t-1}$	-0.00017	0.00025	-0.70	0.49
$AGE_{i,t-1}*CL_{i,t-1}$	0.0005	0.0013	0.43	0.67
$AGE_{i,t-1}*K_{i,t-1}$	-0.0032	0.0013	-2.45	0.01
$AGE_{i,t-1}*I^N_{i,t-1}$	0.0042	0.0031	1.35	0.18
$TL_{i,t-1}*NW_{i,t-1}$	-0.00007	0.00017	-0.41	0.68
$TL_{i,t-1}*CL_{i,t-1}$	0.00077	0.00080	0.97	0.33
$TL_{i,t-1}*K_{i,t-1}$	-0.00147	0.00080	-1.83	0.07
$TL_{i,t-1}*I^N_{i,t-1}$	0.0006	0.0017	0.36	0.72
$NW_{i,t-1}*CL_{i,t-1}$	-0.00015	0.00035	-0.43	0.67
$NW_{i,t-1}*K_{i,t-1}$	0.00103	0.00043	2.38	0.02
$NW_{i,t-1}*I^N_{i,t-1}$	-0.00209	0.00089	-2.34	0.02
$CL_{i,t-1}*K_{i,t-1}$	0.0034	0.0017	1.97	0.05
$CL_{i,t-1}*I^N_{i,t-1}$	0.0016	0.0044	0.37	0.71
$K_{i,t-1}*I^N_{i,t-1}$	0.0118	0.0048	2.45	0.01
Covariance Parameters	Estimate			
σ_y^2	0.29			
σ_ε^2	4.78			

Table 5.2. Mean and median values of the variables

Variable	Mean	Median
$\Delta Q_{i,t}$	10,415.77	8,022.41
$V_{i,t}$	170,280.24	142,557.00
$C_{i,t}$	22.65	22.75
$AGE_{i,t-1}$	46.90	46.00
$TL_{i,t-1}$	170,890.85	120,000.00
$NW_{i,t-1}$	450,020.65	337,062.40
$CL_{i,t-1}$	62,650.95	32,259.41
$K_{i,t-1}$	99,579.27	80,464.00
$I^N_{i,t-1}$	7,103.65	1,385.89
$I^N_{i,t}$	5,001.21	-287.35

Table 5.3. Expected changes in investment and elasticities

Variable	Expected change at the mean	Expected change at the median	Elasticity at the mean	Elasticity at the median
Change in output ($\Delta Q_{i,t}$)	0.039	0.037	0.081	-1.025
Value of short-term assets ($V_{i,t}$)	0.038	0.039	1.284	-19.280
Cost of capital ($C_{i,t}$)	0.058	0.075	2.627	-59.412
Lagged operator age ($AGE_{i,t-1}$)	-0.034	-0.029	-3.207	46.223
Lagged total liabilities ($TL_{i,t-1}$)	0.002	0.001	0.058	-0.463
Lagged net worth ($NW_{i,t-1}$)	0.007	0.009	0.630	-10.443
Lagged current liabilities ($CL_{i,t-1}$)	0.007	0.009	0.088	-0.966
Lagged machinery value ($K_{i,t-1}$)	-0.050	-0.061	-0.996	17.102
Lagged investment ($I^N_{i,t-1}$)	-0.051	-0.075	-0.072	0.362

elasticities. For both the mean and median, the cost of capital has the largest absolute impact on the expected value of investment. The positive impact indicates that as the cost of investment rises, the level of investment also rises. This is the opposite of what was expected; however, if we set all nonsignificant parameter estimates to zero, then the impact of the cost of capital on investment is zero at the mean and slightly above zero at the median. Lagged machinery investment and machinery value have the next largest impacts, as we see a five cent reduction in current investment for each dollar of lagged investment and machinery value. For each dollar change in output or held in short-term assets, investment increases by nearly four cents. As operator age rises by a year, investment decreases by just over three cents. The three financial variables (net worth and the liability variables) have a minimal absolute impact on investment.

The magnitude of elasticities provides evidence of the relative responsiveness of investment to the regressors. At the mean values, operator age has the largest elasticity, followed by cost of capital, the value of short-term assets, lagged machinery value, and net worth. The pattern is nearly the same at the median values, with all elasticity figures increasing in absolute value.

These classical mixed model results are useful as a comparison for the Bayesian results shown later. In one of the Bayesian simulations, we leave the outlier detection component out of the model. This simulation is the most closely aligned to the model above. The other Bayesian simulations extend the model from this point.

5.2 Bayesian Simulation with Variable Selection but no Outlier Detection

For the simulation with variable selection but no outlier detection, we simply remove the outlier detection component from our model and Gibbs sampler. This effectively sets the outlier hyperparameter (η) and the outlier indicators ($\theta_{i,t}$) equal to one. As stated in Section 4.6, the simulations consists of four chains of 12,000 loops, for a total of 48,000 iterations. For the variable selection component, the prior probability for including each of the variables is set at 0.9, implying a 90 percent prior belief that each of the variables belongs in the model. The exceptions to this are for the intercept and the random effects which are always included in the model. Table 5.4 contains the summary statistics for the simulations. The mean values, the sample quantiles, and Gelman and Rubin's R-statistics are computed from the last halves of the chains. The percentage of times the variables are chosen for the model includes information from all 48,000 iterations.

Since $\sqrt{\hat{R}}$ is below 1.03 for all of the quantities of interest, we assume that the chains have converged to their stationary distributions. The variables that were selected for the model over eighty percent of the time are the quadratic effect for the lagged investment

Table 5.4. Summary of the results for the no outlier case

	Posterior	Posterior quantiles				% of times
Variable	Mean	2.5%	50%	97.5%	$\sqrt{\hat{R}}$	chosen
Intercept	0.622	-0.373	0.642	1.544	1.02	100.00
$\Delta Q_{i,t}$	0.030	0.009	0.030	0.049	1.00	97.89
$V_{i,t}$	0.039	0.022	0.038	0.060	1.01	99.99
$C_{i,t}$	0.109	0.000	0.116	0.259	1.00	80.45
$AGE_{i,t-1}$	-0.021	-0.033	-0.020	-0.009	1.00	99.11
$TL_{i,t-1}$	0.000	0.000	0.000	0.002	1.00	4.50
$NW_{i,t-1}$	0.001	0.000	0.000	0.007	1.00	14.54
$CL_{i,t-1}$	0.000	-0.002	0.000	0.007	1.00	7.83
$K_{i,t-1}$	-0.050	-0.073	-0.050	-0.023	1.00	98.79
$I^N_{i,t-1}$	-0.015	-0.090	0.000	0.007	1.00	38.15
$\Delta Q_{i,t}^2$	0.000	0.000	0.000	0.000	1.00	0.33
$V_{i,t}^2$	-0.000	-0.000	0.000	0.000	1.01	11.22
$C_{i,t}^2$	-0.024	-0.085	-0.012	0.000	1.00	54.87
$AGE_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.34
$TL_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.10
$NW_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.10
$CL_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.76
$K_{i,t-1}^2$	-0.000	-0.002	0.000	0.000	1.02	4.00
$I^{N\,2}_{i,t-1}$	0.015	0.000	0.016	0.024	1.00	93.01
$\Delta Q_{i,t}*V_{i,t}$	-0.000	0.000	0.000	0.000	1.00	1.59
$\Delta Q_{i,t}*C_{i,t}$	0.000	0.000	0.000	0.004	1.00	5.43
$\Delta Q_{i,t}*AGE_{i,t-1}$	-0.000	0.000	0.000	0.000	1.00	2.18
$\Delta Q_{i,t}*TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.75
$\Delta Q_{i,t}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.15
$\Delta Q_{i,t}*CL_{i,t-1}$	-0.000	-0.001	0.000	0.000	1.00	5.00
$\Delta Q_{i,t}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.02	1.25
$\Delta Q_{i,t}*I^N_{i,t-1}$	0.000	0.000	0.000	0.003	1.00	4.91
$V_{i,t}*C_{i,t}$	0.000	0.000	0.000	0.006	1.00	8.52
$V_{i,t}*AGE_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.54
$V_{i,t}*TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.44
$V_{i,t}*NW_{i,t-1}$	-0.000	-0.000	0.000	0.000	1.01	22.63
$V_{i,t}*CL_{i,t-1}$	-0.000	-0.000	0.000	0.000	1.01	3.46
$V_{i,t}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.97
$V_{i,t}*I^N_{i,t-1}$	0.000	0.000	0.000	0.005	1.00	8.95
$C_{i,t}*AGE_{i,t-1}$	-0.000	-0.000	0.000	0.000	1.00	3.19
$C_{i,t}*TL_{i,t-1}$	-0.000	-0.003	0.000	0.000	1.00	4.81
$C_{i,t}*NW_{i,t-1}$	0.000	0.000	0.000	0.001	1.01	3.06
$C_{i,t}*CL_{i,t-1}$	-0.000	-0.007	0.000	0.000	1.00	7.96

Table 5.4. (continued)

Variable	Posterior Mean	Posterior quantiles 2.5%	50%	97.5%	$\sqrt{\hat{R}}$	% of times chosen
$C_{i,t}*K_{i,t-1}$	-0.011	-0.023	-0.012	0.000	1.00	77.85
$C_{i,t}*I^N_{i,t-1}$	0.003	-0.001	0.000	0.032	1.00	19.45
$AGE_{i,t-1}*TL_{i,t-1}$	-0.000	-0.001	0.000	0.000	1.00	5.18
$AGE_{i,t-1}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.18
$AGE_{i,t-1}*CL_{i,t-1}$	-0.000	0.000	0.000	0.000	1.00	1.24
$AGE_{i,t-1}*K_{i,t-1}$	-0.000	-0.003	0.000	0.000	1.00	19.62
$AGE_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	2.48
$TL_{i,t-1}*NW_{i,t-1}$	-0.000	-0.000	0.000	0.000	1.00	3.31
$TL_{i,t-1}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.24
$TL_{i,t-1}*K_{i,t-1}$	-0.000	0.000	0.000	0.000	1.00	1.07
$TL_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.003	1.00	9.51
$NW_{i,t-1}*CL_{i,t-1}$	-0.000	-0.001	0.000	0.000	1.00	21.33
$NW_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.26
$NW_{i,t-1}*I^N_{i,t-1}$	-0.000	0.000	0.000	0.000	1.00	1.39
$CL_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	1.14
$CL_{i,t-1}*I^N_{i,t-1}$	0.003	0.000	0.000	0.010	1.00	45.88
$K_{i,t-1}*I^N_{i,t-1}$	0.003	0.000	0.000	0.014	1.00	30.09
Parameter						
σ_ε^2	4.912	4.591	4.908	5.251	1.00	
σ_γ^2	0.498	0.266	0.465	0.928	1.00	

level and the linear effects for the change in output, the value of short-term assets, the cost of capital, lagged operator age, and lagged machinery value. This list differs from the list of variables with statistically significant parameter values from the classical mixed model. While the classical model indicated seven of the cross effects had a significant impact on investment, no cross effects were chosen by the variable selection procedure employed here.

To examine the differences between the classical mixed model results and the results given in Table 5.4, let us look at the fifteen variables that either had significant parameter values from the classical mixed model or had an eighty percent inclusion rate in the Bayesian model. For ten of those fifteen variables, the classical mixed model parameter estimate fell within the 95 percent sample quantile range from the Bayesian results. In four of these cases, the variable was chosen by the Bayesian model less than one-quarter of the time. This means that over 75 percent of the sample for that parameter is set at zero and the sample quantile range is likely to be very small. Lagged operator age is the only variable that was selected a vast majority of times (99.11 percent), but whose classical mixed model parameter estimate did not fall with the sample quantile range.

For the variance parameter estimates, both the error variance and the random effects variance estimates from the classical mixed model fall within the sample quantile ranges. The mean estimate for the error variance is greater than that from the classical mixed model, 4.91 versus 4.78. The mean estimate for the random effects variance from this procedure is greater than the estimate from the classical mixed model, 0.50 versus 0.29. Given the informative prior for the random effects variance and the small size of the time series in the panel data set, the draws for the random effects variance show significant influence from the prior distribution.

The classical mixed model results indicated that four of the five variables inserted to cover the 5 Cs of the farmer credit situation had a significant impact on the farmer's investment decision. The Bayesian model with no outliers only supports the inclusion of lagged operator age and the value of short-term assets. The significant effects from the classical mixed model that included net worth and/or current liabilities were selected less than half of the time.

To provide more detail on the movement of the chains through the parameter space, we have graphed the chains for the parameters that were selected at least eighty percent of the time and for the variance parameters. Figures 5.1 through 5.8 display these graphs. The graphs show that, at least for these variables, the behavior of the chains converged at a rapid rate. During most of the Bayesian simulations in this study, the intercept and the random effect variables were the last to converge. Figures 5.9 and 5.10 show chain behavior for variables that were selected less often. In Figure 5.9, the chain behavior for the parameter for the cross effect $CL_{i,t-1}{*}I^{N}_{i,t-1}$ is displayed. This effect was chosen in nearly 46 percent of the trials. The cross effect $\Delta Q_{i,t}{*}NW_{i,t-1}$ was chosen less than one percent of the time. The chain behavior for its parameter is shown in Figure 5.10. These graphs show the impact of not being chosen for the model through the solid mass at zero. Overall, the ten figures show that the chains moved quickly to the target stationary distributions.

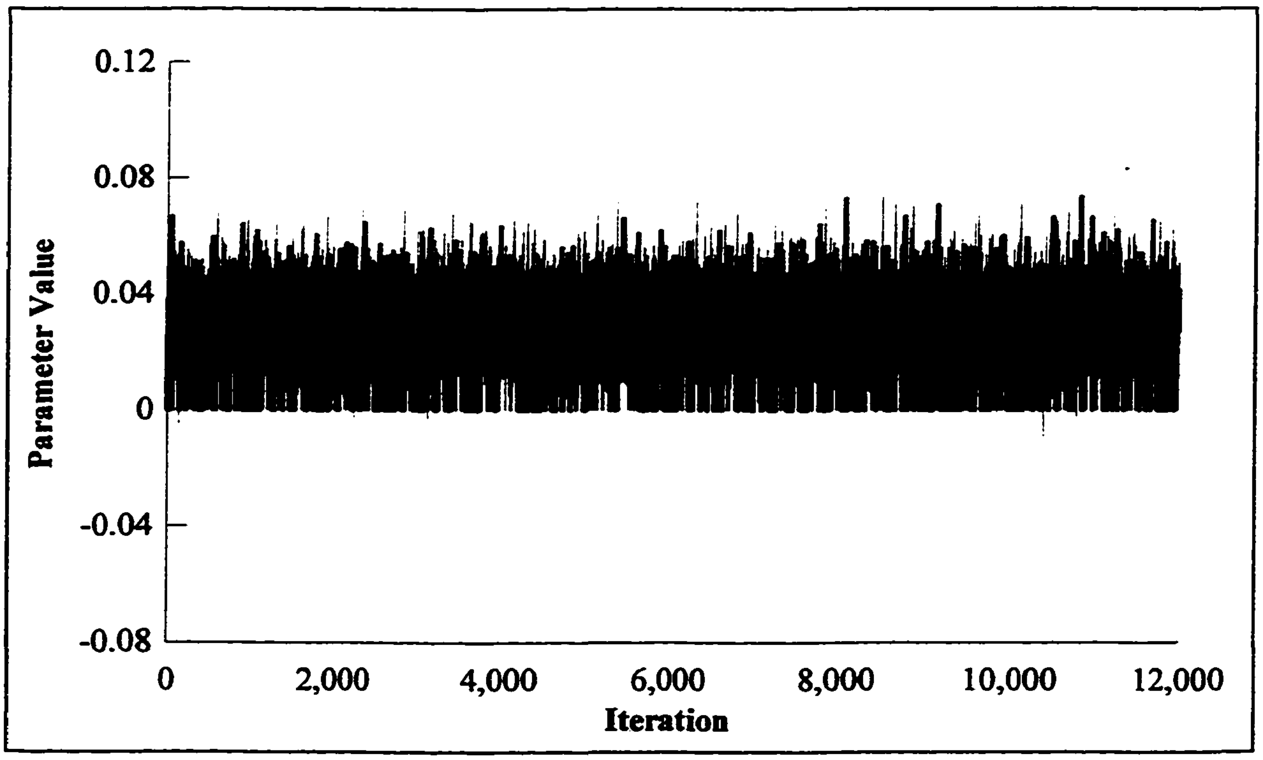


Figure 5.1. Graph of the chains for β_1, the parameter for $\Delta Q_{i,t}$

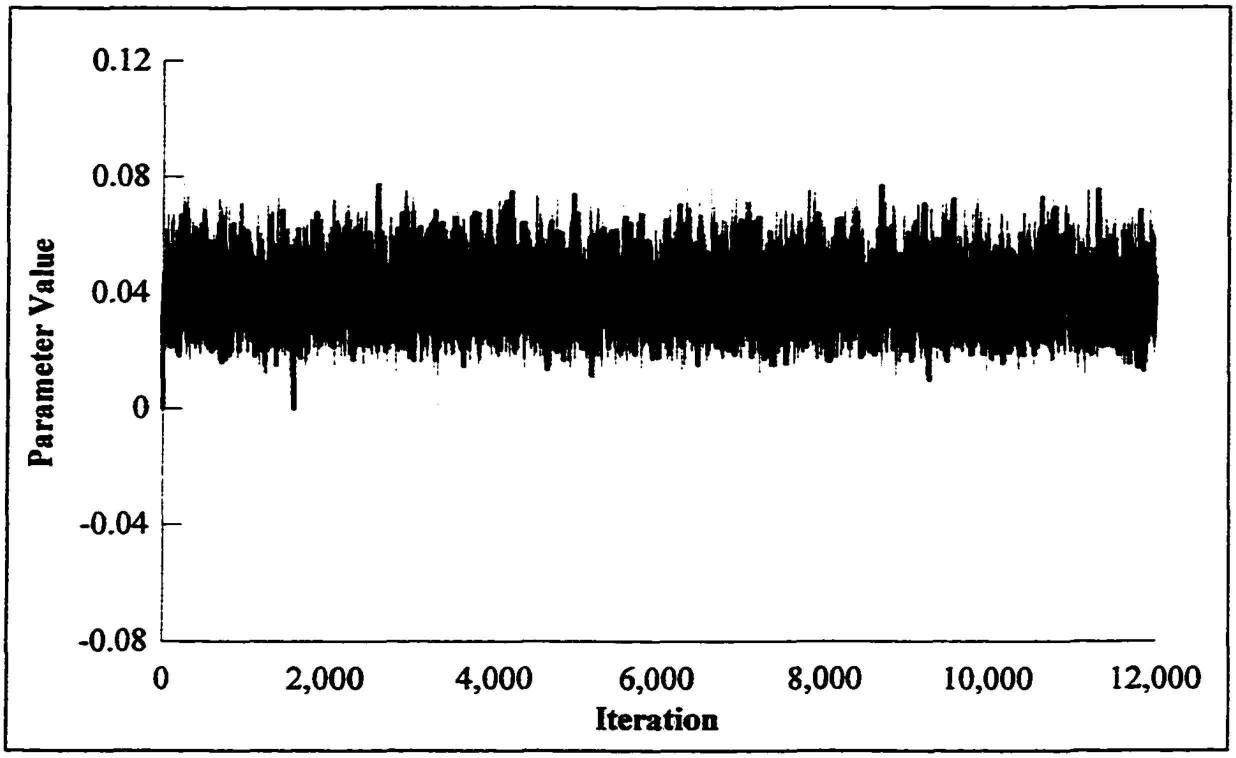


Figure 5.2. Graph of the chains for β_2, the parameter for $V_{i,t}$

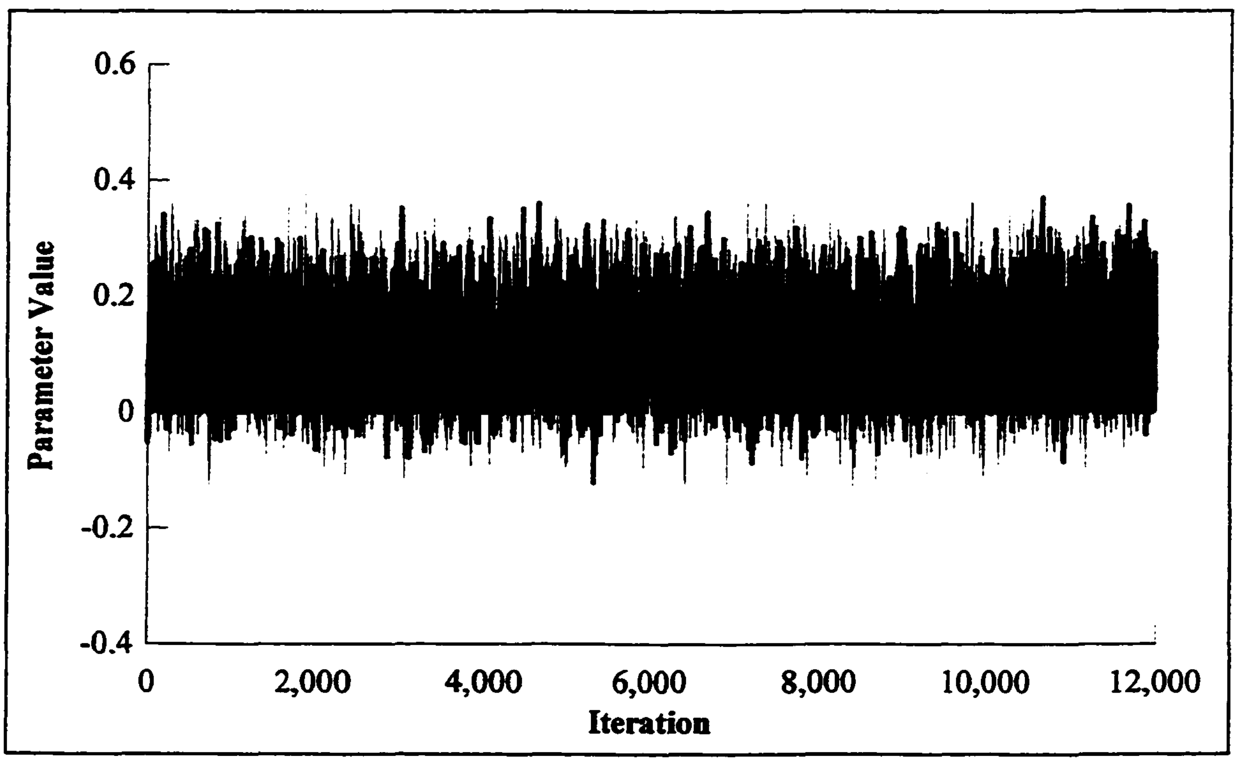


Figure 5.3. Graph of the chains for β_3, the parameter for $C_{i,t}$

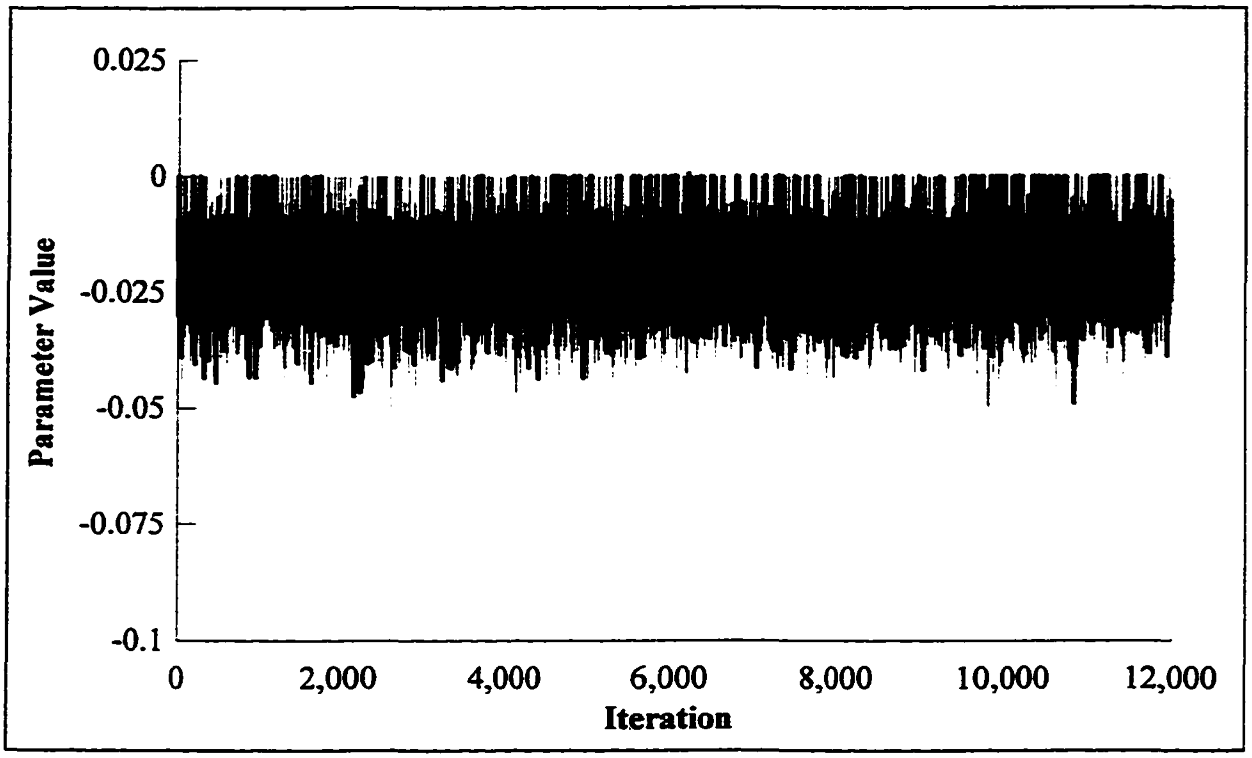


Figure 5.4. Graph of the chains for β_4, the parameter for $AGE_{i,t-1}$

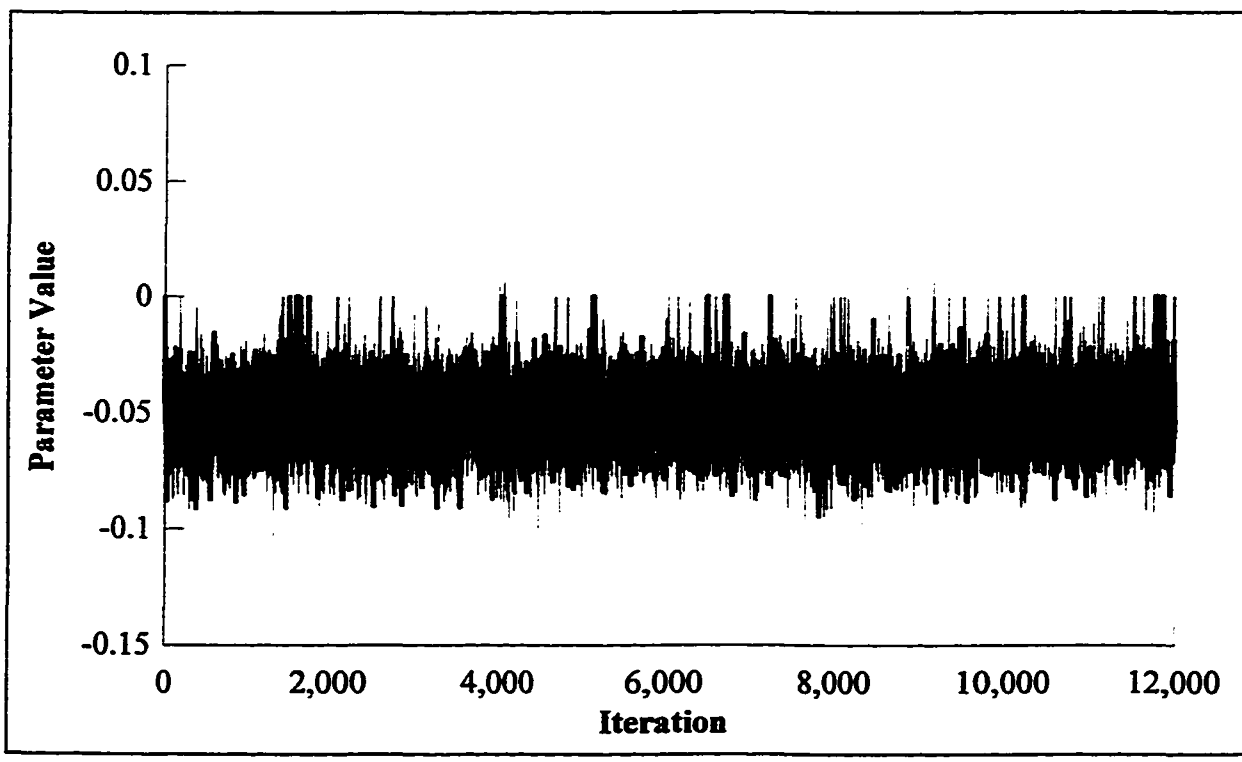


Figure 5.5. Graph of the chains for β_8, the parameter for $K_{i,t-1}$

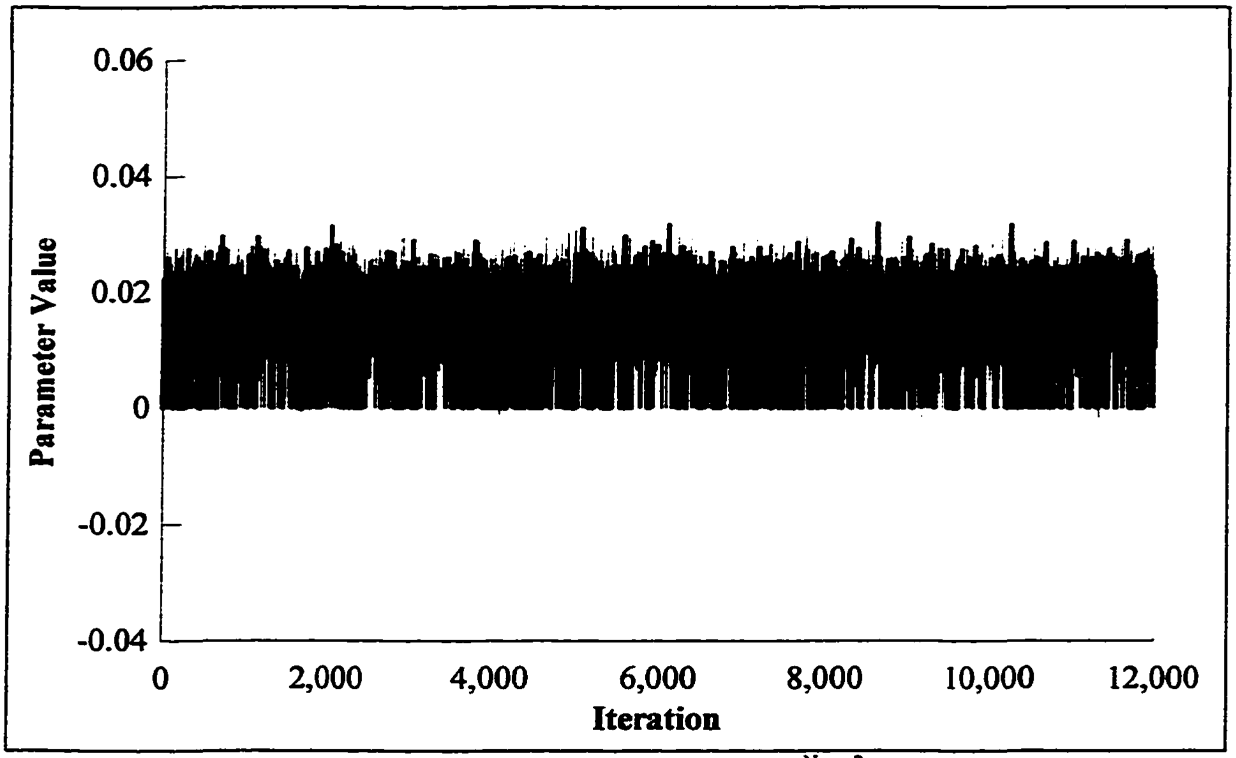


Figure 5.6. Graph of the chains for β_{18}, the parameter for ${I^{N}_{i,t-1}}^2$

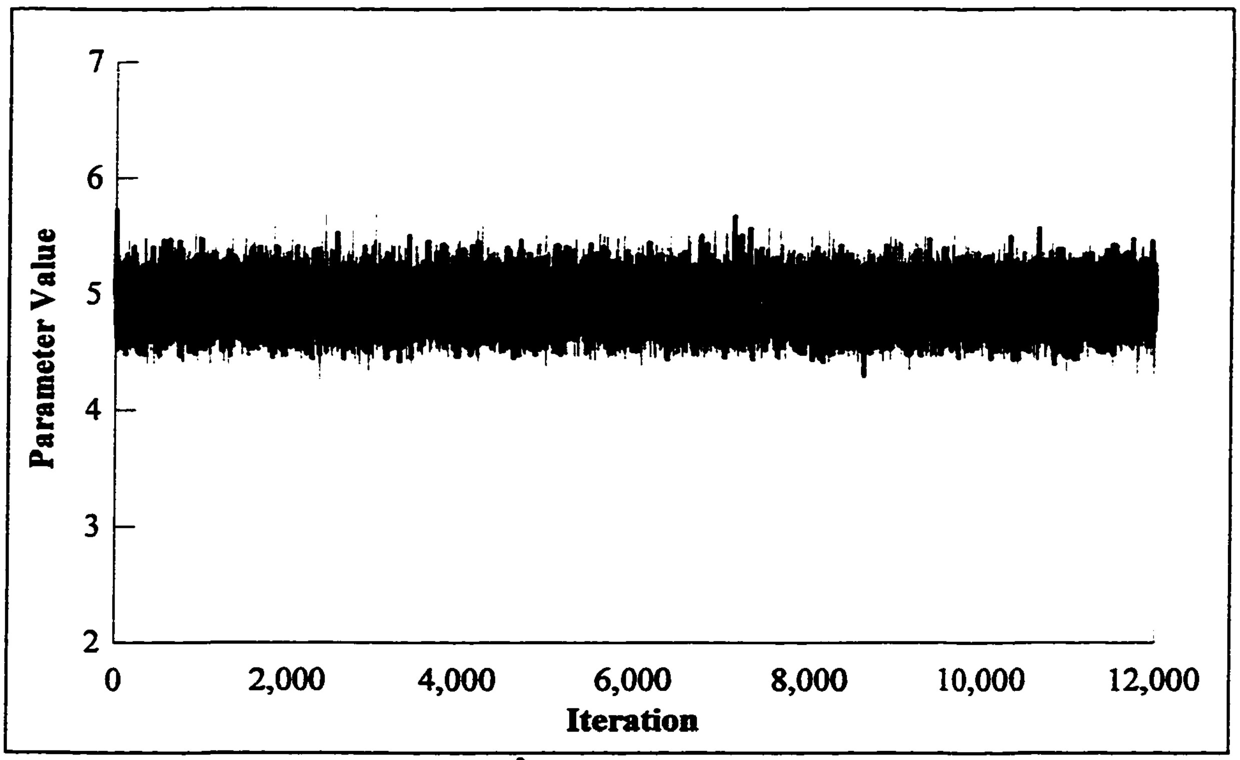


Figure 5.7. Graph of the chains for σ_ε^2, the error variance parameter

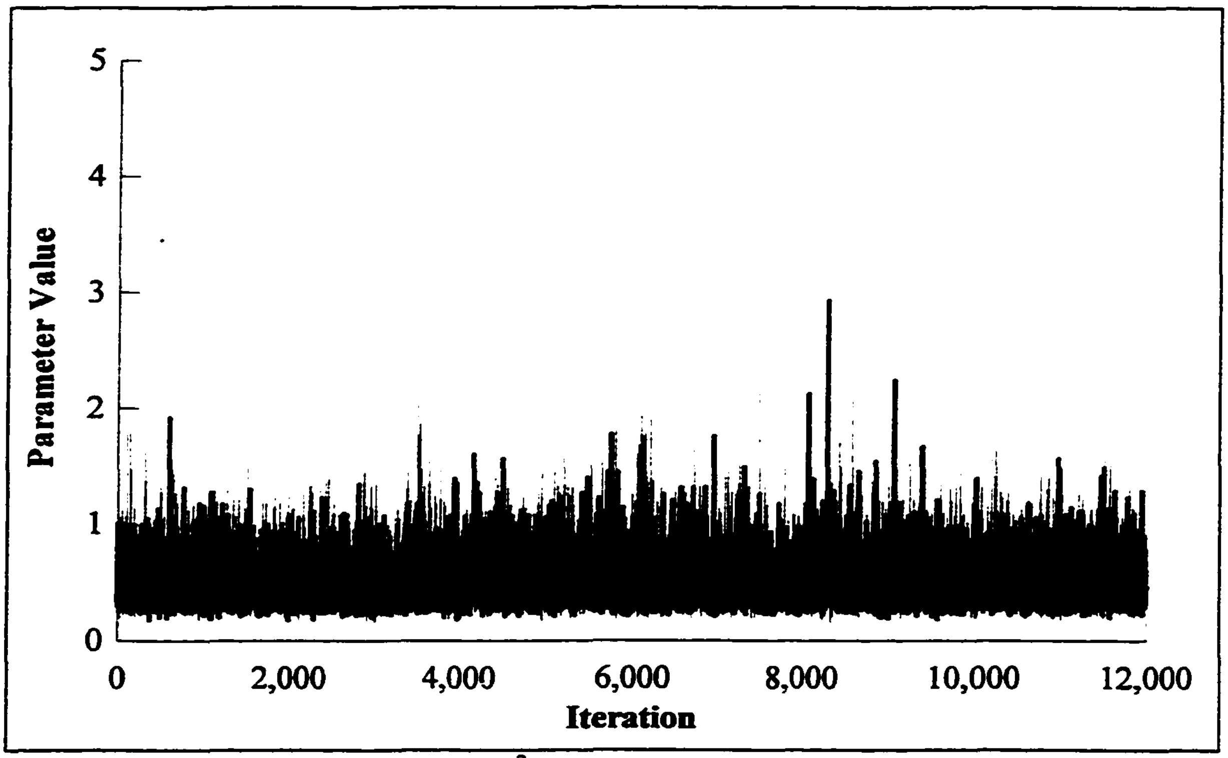


Figure 5.8. Graph of the chains for σ_y^2, the random effects variance parameter

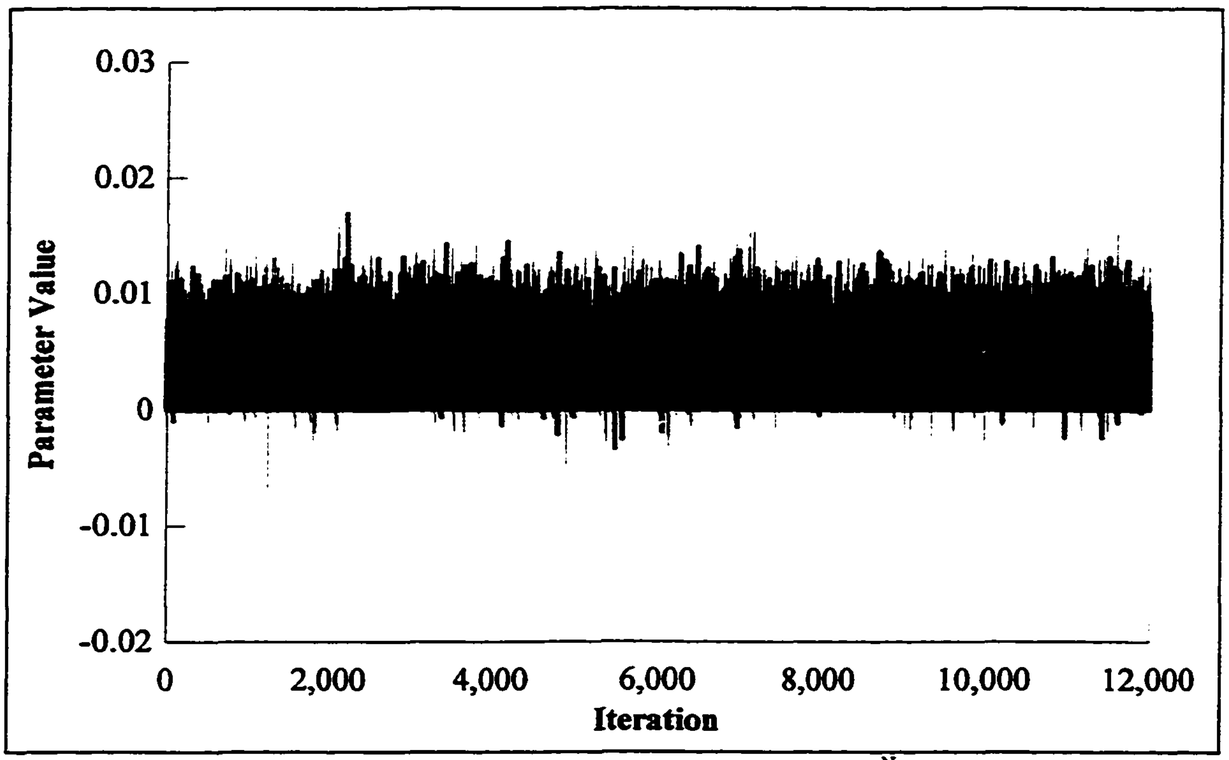


Figure 5.9. Graph of the chains for β_{53}, the parameter for $CL_{i,t-1}*I^N_{i,t-1}$

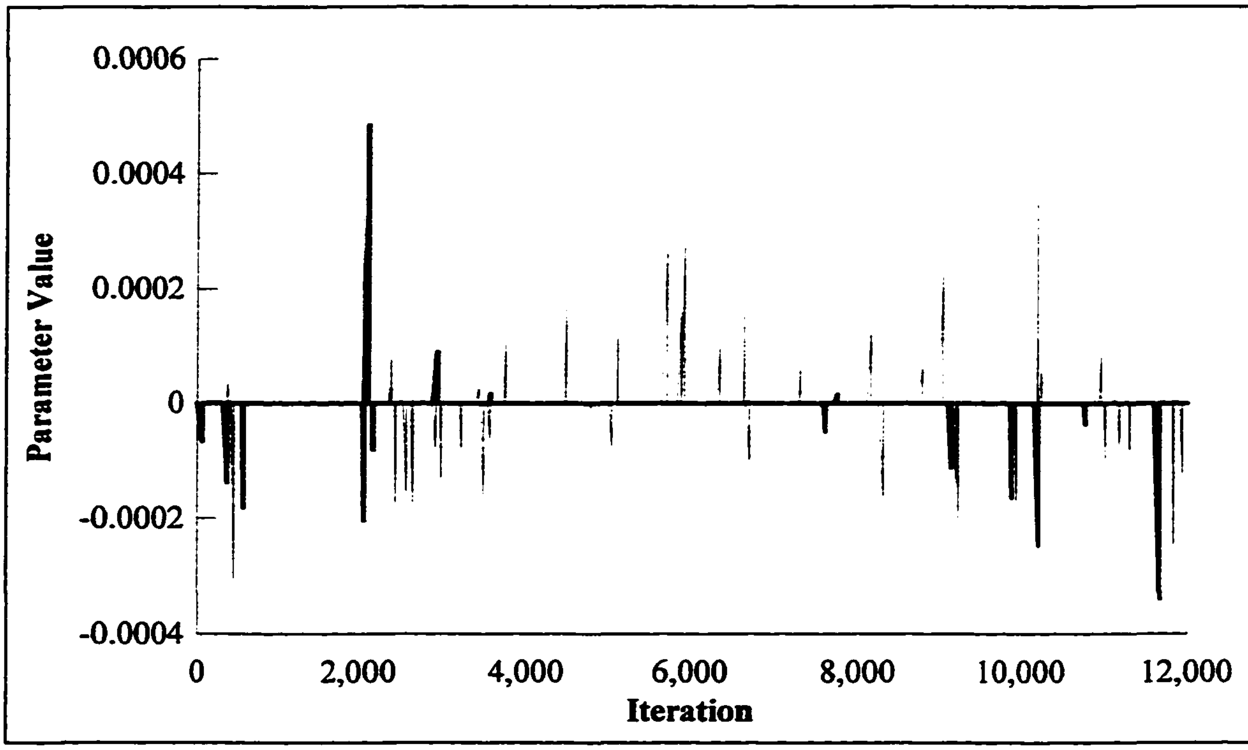


Figure 5.10. Graph of the chains for β_{23}, the parameter for $\Delta Q_{i,t}*NW_{i,t-1}$

5.3 Bayesian Simulation with Variable Selection and Outlier Detection

The only changes to the model for this section from the model for the previous section is the addition of the outlier detection component and the change in the prior probability values for the variable selection component. We have set the prior probabilities for variable inclusion at 0.5 and the priors for the outlier detection component indicate that ten percent of the observations are outliers. We use this scenario as our "base" scenario for the Bayesian estimations with both the variable selection and outlier detection. In the next section, we vary these prior settings and compare the results to those in the present section.

The estimation consists of four chains with 12,000 loops, for a total of 48,000 iterations. The summary statistics for the estimation are given in Table 5.5. As with the no outlier scenario, the mean parameter values, sample quantiles, and the R-statistics are computed from the later halves of the chains and the variable inclusion percentages count all iterations. This same format is employed in summarizing all of the Bayesian estimation scenarios.

With $\sqrt{\hat{R}}$ below 1.2 for almost all of the parameters, the model is considered converged. The exception is the parameter for $K_{i,t-1}*I^N_{i,t-1}$, which was selected for the model less than two percent of the time; thus, we chose not to extend the chain length and re-estimate. The variable selection component chose four of the fifty-four variables over eighty percent of the iterations. These variables are the change in output, the value of short-term assets, lagged machinery value, and the square of lagged investment. The outlier detection hyperparameter (η) reached an average value of 0.74, indicating a posterior proportion of 0.74 that an observation is not a potential outlier. The error and random effects variance average estimates are 0.922 and 0.470, respectively. The error variance estimate is much smaller for this formulation due to the outlier detection component and the variance inflation parameter, κ. The classical mixed model and the no outlier Bayesian formulation have to accommodate any outliers by increasing the size of the error variance estimate.

The average parameter values for the four selected variables indicate that machinery investment rises with increases in changes in output, short-term asset value, and the square of lagged investment and declines with increases in lagged machinery value. These relationships are the same indicated through the Bayesian no outlier scenario and the classical mixed model results.

Table 5.5. Summary of the results for the full model

Variable	Posterior Mean	Posterior quantiles 2.5%	50%	97.5%	$\sqrt{\hat{R}}$	% of times chosen
Intercept	-0.251	-1.166	-0.246	0.671	1.08	100.00
$\Delta Q_{i,t}$	0.023	0.000	0.024	0.043	1.00	88.33
$V_{i,t}$	0.017	0.010	0.017	0.025	1.00	99.59
$C_{i,t}$	0.007	0.000	0.000	0.089	1.01	13.70
$AGE_{i,t-1}$	-0.004	-0.013	0.000	0.000	1.00	47.81
$TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.28
$NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.19
$CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.01	0.50
$K_{i,t-1}$	-0.060	-0.073	-0.060	-0.046	1.00	99.99
$I^N_{i,t-1}$	-0.040	-0.091	-0.047	0.000	1.00	69.05
$\Delta Q_{i,t}^2$	0.000	0.000	0.000	0.000	1.00	0.09
$V_{i,t}^2$	0.000	0.000	0.000	0.000	1.00	0.46
$C_{i,t}^2$	-0.000	0.000	0.000	0.000	1.01	2.39
$AGE_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.04
$TL_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.02
$NW_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.01
$CL_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.10
$K_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.04
$I^N_{i,t-1}{}^2$	0.020	0.011	0.020	0.026	1.04	97.70
$\Delta Q_{i,t}*V_{i,t}$	0.000	0.000	0.000	0.000	1.00	0.13
$\Delta Q_{i,t}*C_{i,t}$	0.000	0.000	0.000	0.000	1.01	0.28
$\Delta Q_{i,t}*AGE_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.06
$\Delta Q_{i,t}*TL_{i,t-1}$	-0.000	-0.001	0.000	0.000	1.00	27.09
$\Delta Q_{i,t}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.03
$\Delta Q_{i,t}*CL_{i,t-1}$	-0.000	-0.001	0.000	0.000	1.02	12.79
$\Delta Q_{i,t}*K_{i,t-1}$	-0.000	0.000	0.000	0.000	1.01	1.71
$\Delta Q_{i,t}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.03	0.33
$V_{i,t}*C_{i,t}$	0.000	0.000	0.000	0.000	1.00	0.23
$V_{i,t}*AGE_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.04
$V_{i,t}*TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.05
$V_{i,t}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	1.07
$V_{i,t}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.38
$V_{i,t}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.09
$V_{i,t}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.04	0.27
$C_{i,t}*AGE_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.20
$C_{i,t}*TL_{i,t-1}$	-0.000	-0.002	0.000	0.000	1.00	3.63
$C_{i,t}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.07
$C_{i,t}*CL_{i,t-1}$	-0.000	0.000	0.000	0.000	1.00	1.99

Table 5.5. (continued)

Variable	Posterior Mean	Posterior quantiles 2.5%	50%	97.5%	$\sqrt{\hat{R}}$	% of times chosen
$C_{i,t}*K_{i,t-1}$	-0.002	-0.012	0.000	0.000	1.00	21.28
$C_{i,t}*I^N_{i,t-1}$	-0.000	0.000	0.000	0.000	1.00	2.54
$AGE_{i,t-1}*TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.03
$AGE_{i,t-1}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.02
$AGE_{i,t-1}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.08
$AGE_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.11
$AGE_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.71
$TL_{i,t-1}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.05
$TL_{i,t-1}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.03
$TL_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.05
$TL_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.10
$NW_{i,t-1}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.15
$NW_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$NW_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.08
$CL_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.09
$CL_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.10	2.05
$K_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.28	1.49
Parameter						
η	0.740	0.702	0.740	0.777	1.00	
σ_ε^2	0.922	0.815	0.920	1.040	1.00	
σ_v^2	0.470	0.251	0.438	0.876	1.01	

In comparing the results from this estimation and the no outlier scenario, we see that the four variables selected here were also selected in the no outlier scenario. The average parameter values for these variables are of the same sign and magnitude. The greatest difference between the two estimates is in the parameter for the value of short-term assets; the mean estimate from the no outlier case is twice the size of the mean estimate from the present case. Two other variables were chosen in the no outlier scenario but not selected here, the cost of capital and lagged operator age. The random effects variance estimates are very similar. The error variance estimates are quite different as would be expected from the addition of the outlier detection component. The error variance mean estimate from the no outlier scenario is five times the mean estimate produced in this formulation.

Table 5.6 shows the breakdown of the percentages of outlier detection for this estimation. Of the 1770 observations, 805 (45.5 percent) were not selected as outliers over 90 percent of the iterations. 1378 observations (77.9 percent) were not selected as outliers over 70 percent of the time, but 207 observations (11.7 percent) were selected as outliers over 90 percent of the time.

Table 5.6. Outlier detection percentages

Percentage range	Number of observations
0 - 10	805
10 - 20	470
20 - 30	103
30 - 40	50
40 - 50	36
50 - 60	29
60 - 70	21
70 - 80	28
80 - 90	21
90 - 100	207

Figures 5.11-5.17 show the chain paths for the parameters of the four selected variables, the variance components, and the outlier detection hyperparameter. As with the no outlier scenario, we can see that the chains "converge" rather quickly. The first six of these figures have counterparts from the no outlier scenario and, for the most part, they have very similar features. The differences can be attributed to differences in the proportion of selection and the mean level of parameter estimates. For example, between Figures 5.2 and 5.12, the graphs of the chains for the parameter of short-term assets, $V_{i,t}$, the only noticeable difference

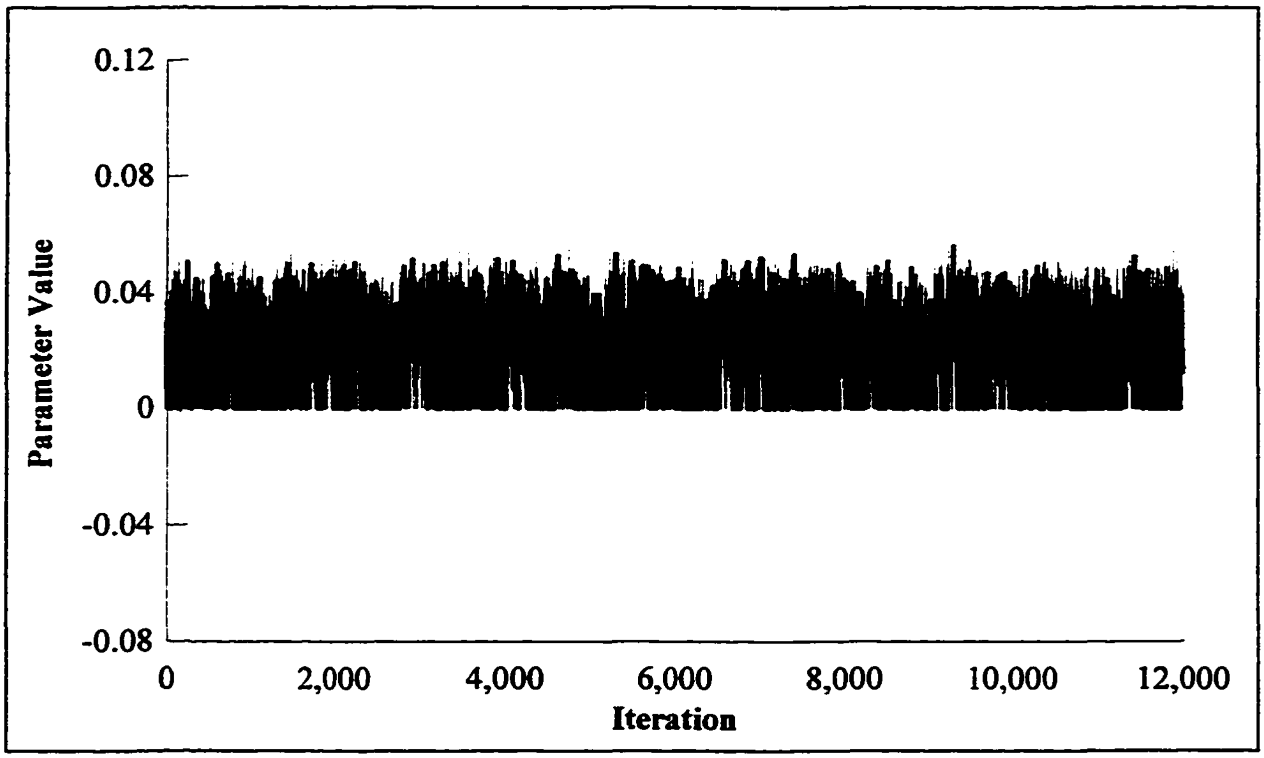


Figure 5.11. Graph of the chains for β_1, the parameter for $\Delta Q_{i,t}$

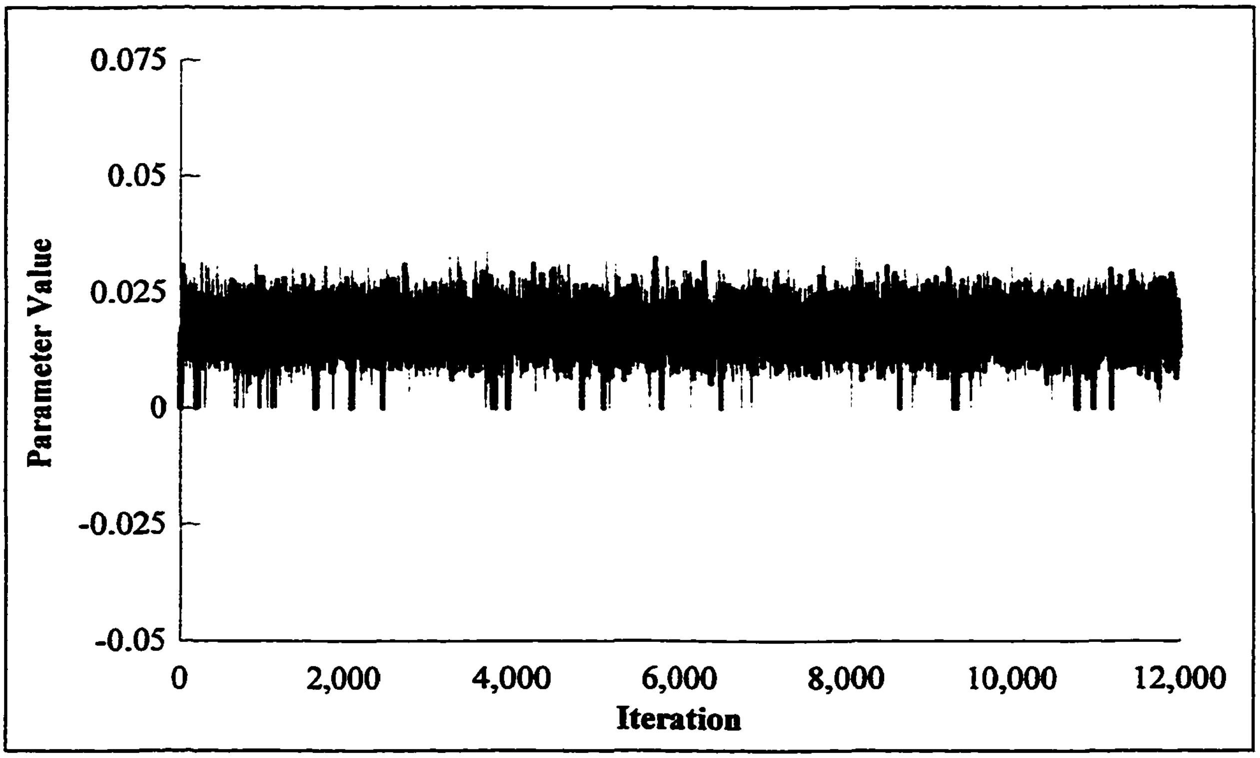


Figure 5.12. Graph of the chains for β_2, the parameter for $V_{i,t}$

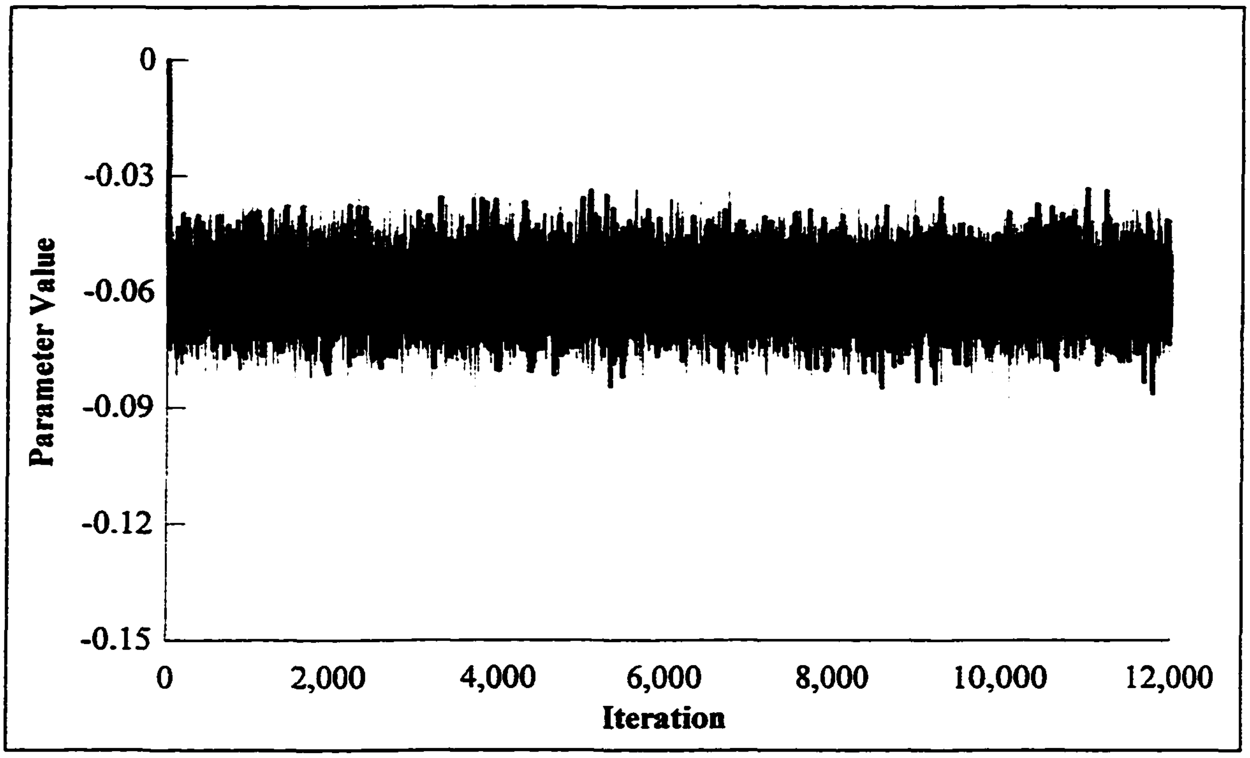


Figure 5.13. Graph of the chains for β_8, the parameter for $K_{i,t-1}$

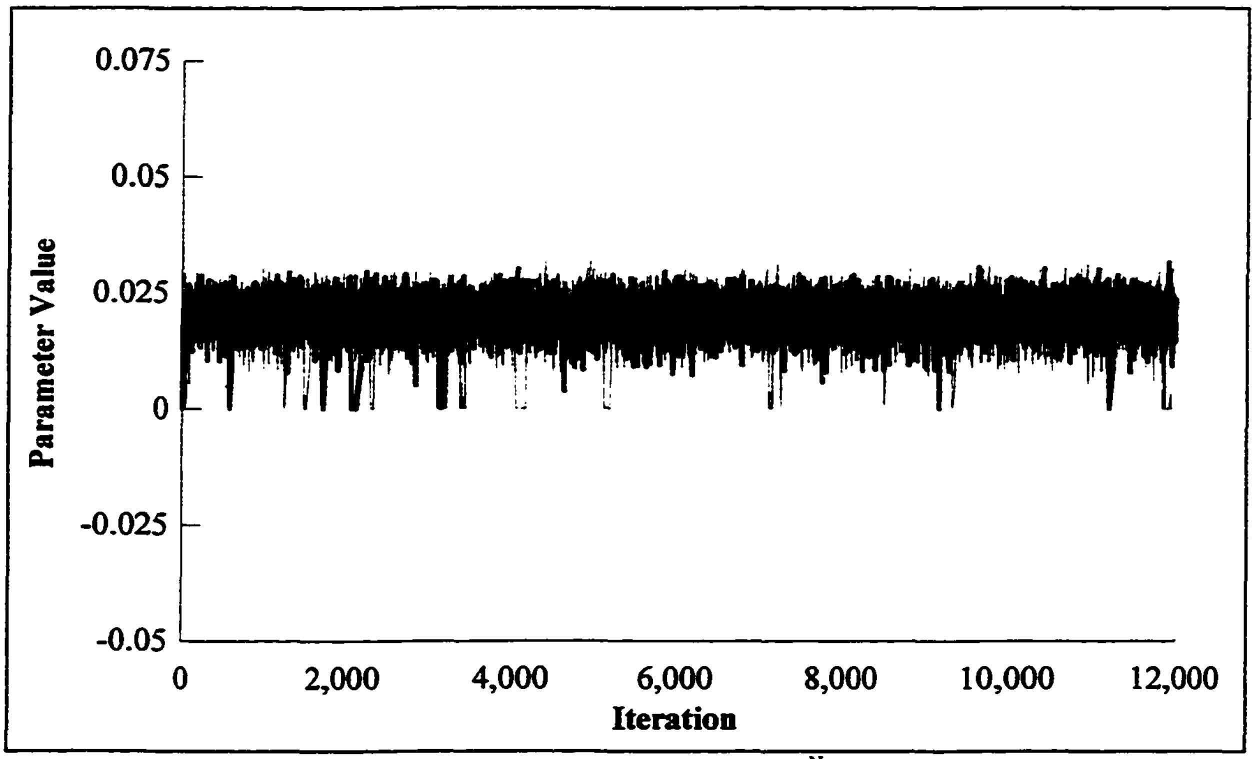


Figure 5.14. Graph of the chains for β_{18}, the parameter for $I^N_{i,t-1}$

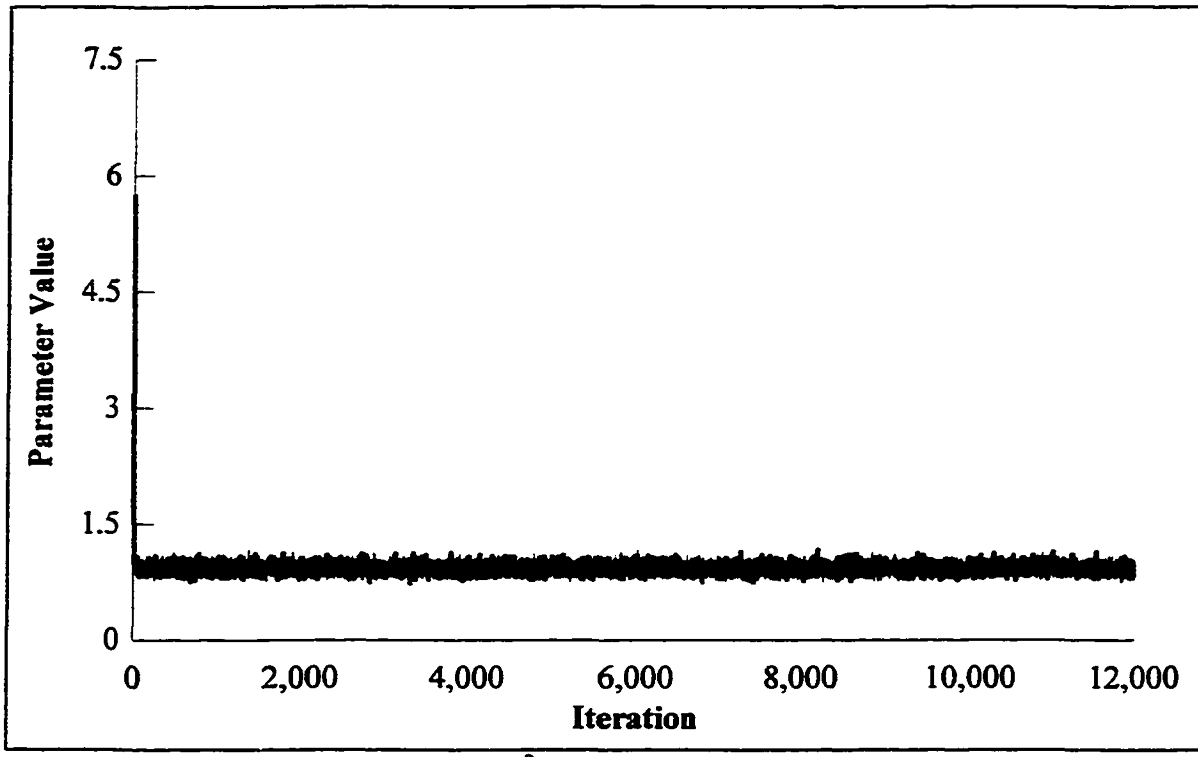


Figure 5.15. Graph of the chains for σ_ε^2, the error variance parameter

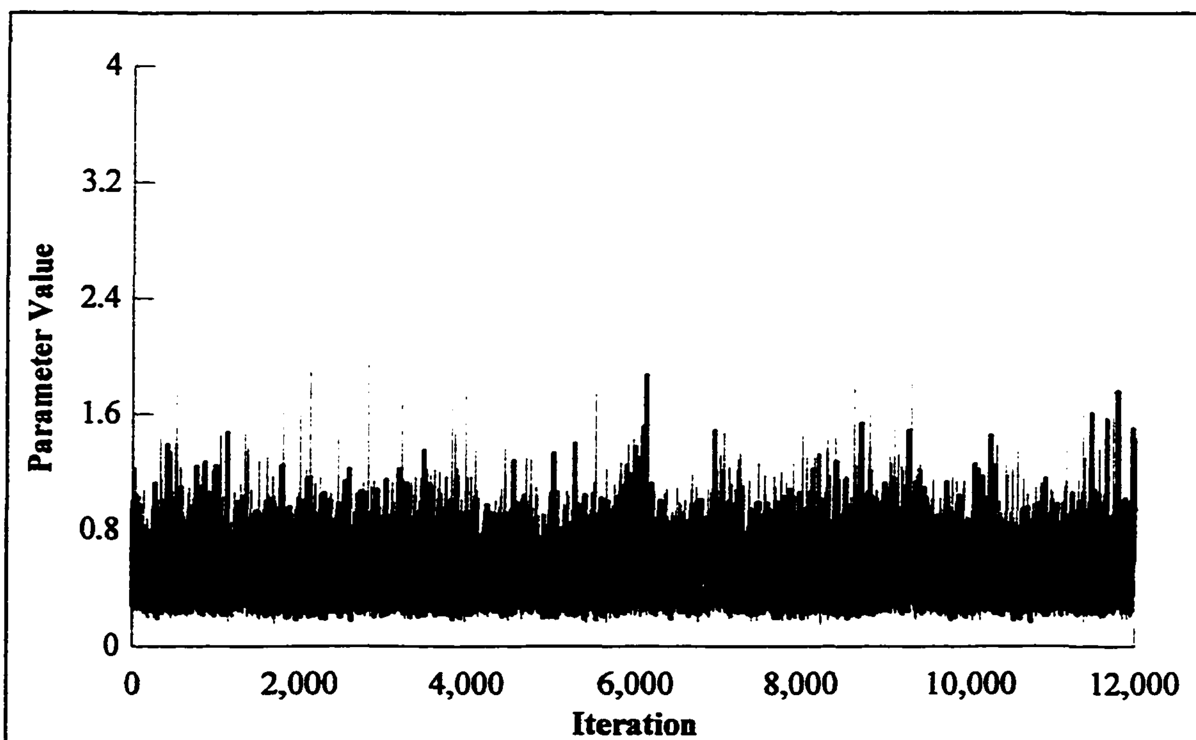


Figure 5.16. Graph of the chains for σ_y^2, the random effects variance parameter

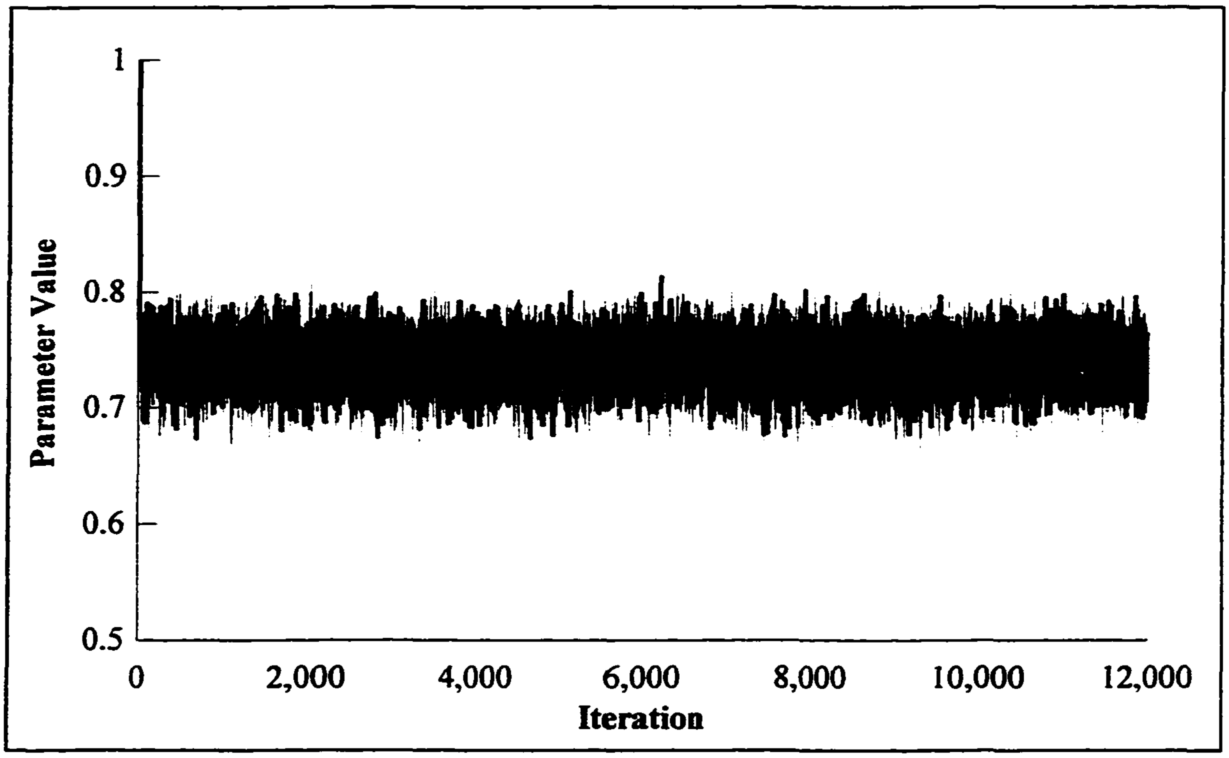


Figure 5.17. Graph of the chains for η, the outlier detection hyperparameter

is the average value of the chain paths. The only truly new figure is Figure 5.17, the graph of the outlier detection hyperparameter chains. The chains settle quickly between 0.65 and 0.8, implying that the percentage of outliers in the data set is between 20 and 35 percent.

5.4 Sensitivity to the Prior Distributions

To test the impacts of the priors for the variable selection and outlier detection components, we formulated the model under eight other prior specifications. The priors combined three settings each for the variable selection and outlier detection components. Table 5.7 shows the prior combinations and provides the names with which we will refer to each Bayesian estimation. The variable selection prior refers to the prior probability that the

Table 5.7. The various prior specifications

Variable Selection	Outlier Detection 0.0	0.1	0.5	0.9
0.1		Var1Out1	Var1Out5	Var1Out9
0.5		Main	Var5Out5	Var5Out9
0.9	NoOut	Var9Out1	Var9Out5	Var9Out9

variables are included in the model. The outlier detection prior refers to the setting of the outlier detection hyperparameters and the prior proportion of outliers they indicate.

The names were chosen to be descriptive of the prior settings. The Bayesian estimation labeled Main is the one described in the previous section and is the base for comparison for the rest of the estimations. We will also refer to the no outlier Bayesian estimation; we have labeled it NoOut.

As with the other Bayesian estimations, the Gibbs sampler is run over four chains of 12,000 loops each for a total of 48,000 iterations. The behavior of the chains was monitored with Gelman and Rubin's R-statistic, $\sqrt{\hat{R}}$. For all of the variables selected for the model in at least three percent of the iterations, $\sqrt{\hat{R}}$ is less than 1.15 in all of the estimations, so convergence is assumed.

A summary table for each of the Bayesian estimations is provided in Appendix 2, but to quickly summarize the results of the estimations and to allow easier comparison between them, we have constructed a composite summary table, Table 5.8. It lists, for each specification, the variables selected at least eighty percent of the iterations, the mean values of their parameters, the mean values of the variance components and outlier detection hyperparameter, and the histogram of outlier detection.

Table 5.8. Composite summary table

	Bayesian estimation									
	Var1 Out1	Var1 Out5	Var1 Out9	Main	Var5 Out5	Var5 Out9	Var9 Out1	Var9 Out5	Var9 Out9	No Out
Variable	Posterior mean parameter value*									
$\Delta Q_{i,t}$				0.023	0.025	0.025	0.028	0.028	0.029	0.030
$V_{i,t}$	0.017	0.017	0.018	0.017	0.017	0.018	0.017	0.017	0.017	0.039
$C_{i,t}$										0.109
$AGE_{i,t-1}$							-0.007	-0.007	-0.007	-0.021
$K_{i,t-1}$	-0.061	-0.060	-0.061	-0.060	-0.060	-0.060	-0.058	-0.059	-0.059	-0.050
$I^N_{i,t-1}$							-0.054	-0.054	-0.055	
${I^N_{i,t-1}}^2$	0.017	0.016	0.015	0.020	0.019	0.019	0.020	0.020	0.020	0.015
Parameter										
σ_ε^2	0.940	0.921	0.879	0.922	0.906	0.866	0.919	0.905	0.863	4.912
σ_y^2	0.458	0.469	0.455	0.470	0.458	0.465	0.461	0.478	0.476	0.498
η	0.741	0.732	0.709	0.740	0.732	0.708	0.742	0.734	0.710	
Outlier %**	Number of observations									
0 - 10	807	734	399	805	730	412	818	756	451	
10 - 20	464	527	816	470	534	804	462	512	770	
20 - 30	103	105	122	103	104	120	108	115	126	
30 - 40	49	47	60	50	49	64	38	40	58	
40 - 50	45	48	42	36	39	36	34	34	35	
50 - 60	23	30	32	29	29	32	33	33	29	
60 - 70	24	20	23	21	22	32	21	21	27	
70 - 80	25	22	28	28	29	23	30	28	29	
80 - 90	29	34	34	21	25	32	21	22	31	
90 - 100	201	203	214	207	209	215	205	209	214	

*If there is no value in a cell, either the variable was not selected eighty percent of the time or the parameter was not estimated in that scenario.

**The percentage of times the observation was chosen as an outlier.

Several definite patterns can be seen in the table. When we examine the subset of variables selected in all of the Bayesian estimations with both variable selection and outlier detection, the posterior mean parameter estimates are very similar. Given a variable selection prior probability, the estimation procedure selects the same set of variables regardless of the prior chosen for the proportion of outliers. Given a variable inclusion prior of ten percent, the estimation method chose the value of short-term assets, lagged machinery value, and squared lagged investment. Raising the variable inclusion prior to fifty percent added change in output to this list. When the prior is at ninety percent, lagged operator age and lagged investment also are selected in at least eighty percent of the model draws. The variance and outlier hyperparameter estimates are also consistent across the estimations.

A similar pattern emerges from the outlier detection results. Selected variables are estimated to have approximately the same values given a prior value for the outlier detection component, regardless of the prior chosen for variable inclusion. Given a prior for the proportion of outliers of ninety percent, the posterior probability for the proportion of outliers is 71 percent. This posterior probability does not change noticeably even when we change the prior proportion of outliers from 10 percent to 90 percent. As we lower the prior probability of outliers, the outlier hyperparameter indicates a slightly smaller percentage of outliers and the largest group on the outlier percentage scale has a zero to ten percent probability of being an outlier.

Given these results, we can state that the choice of priors for the variable selection and outlier detection components have a negligible effect on the results from the Bayesian analysis. However, the addition of an outlier detection component has several impacts on the results, when compared to the no outlier scenario. For these comparisons, we concentrate on the Bayesian results with the ninety percent prior probability for variable inclusion since the no outlier scenario was also run using that prior. The no outlier case excluded lagged investment in favor of cost of capital. The posterior means for value of short-term assets and lagged operator age more than doubled under the no outlier scenario. The posterior mean of the error variance is five times greater under the no outlier case versus the scenarios including outliers. This is to be expected, because the error variance must be large enough to accommodate all the outliers that are not classified as such. Even with these differences, the no outlier and outlier detection scenarios held several similarities. The posterior mean for change in output is nearly identical in both cases. The point estimates for the parameters associated with lagged machinery value and the square of lagged investment are contained within the posterior quantile ranges (2.5 to 97.5 percent) for the opposite case. The posterior means of the random effects variance are quite similar.

5.5 Bayesian Elasticity Estimates

We estimated the marginal posterior distributions of the expected change in investment related to a one unit change in each regressor and of the elasticity $\left(\frac{\partial y}{\partial x}\frac{x}{y}\right)$ for each regressor, to examine the absolute and relative impacts of the factors on farm machinery investment. As before, we compute the posterior distributions of these measures at both the mean and median values for all variables because of the skewness in the investment data. The MCMC approach allows us to approximate the marginal posterior distributions of unit changes and elasticities since the latter portions of the chains can be thought of as coming from the posterior distributions of interest. As we mentioned in Section 3.1, the MCMC approach provides a simple mechanism for approximating posterior distributions of any measurable function of the model parameters. Expected changes and elasticities are functions of the parameters of our model. Within the Bayesian framework, we then obtain point estimates and credible intervals (the Bayesian equivalent to frequentist confidence intervals) for the expected changes and elasticities. This approach also allows us to incorporate the uncertainty about all model parameters directly into the expected change and elasticity estimates. From the parameter draws from the Gibbs sampler for the Main case (which had prior expectations of 50 percent for the variable inclusion component and of 10 percent for the outlier detection component), we computed marginal posterior distributions of the elasticity measures summarized in Table 5.9. Again, since current investment is positive at the mean and negative at the median, the elasticity measures change signs.

The table shows the posterior means, selected posterior quantiles, Gelman and Rubin's R-statistics, and the percentage of times the estimate is non-zero (out of 48,000 iterations). We see that at the mean values lagged machinery value has the largest relative impact followed by the value of short-term assets, lagged operator age, and the cost of capital. The financial variables representing the 5 Cs of lending, lagged net worth, total liabilities, and current liabilities, have a negligible impact. Non-zero estimates for these elasticities only occur less than 0.5 percent of the time. At the median values, the order is nearly the same but with the cost of capital becoming more influential.

Given the posterior means of the elasticities, we calculated the expected changes in investment given in Table 5.10. At the mean values, lagged machinery value has the largest absolute impact on investment, followed by lagged investment, changes in output, and the value of short-term assets. At the median values, only lagged machinery value and lagged investment change places. At both levels, the lagged financial variables have again negligible

Table 5.9. Summary of the elasticity results for Main

Variable	Posterior Mean	Posterior quantiles 2.5%	50%	97.5%	$\sqrt{\hat{R}}$	% of times non-zero
at the mean						
$\Delta Q_{i,t}$	0.048	0.000	0.050	0.089	1.00	88.33
$V_{i,t}$	0.585	0.332	0.583	0.855	1.00	99.59
$C_{i,t}$	0.333	0.000	0.000	4.027	1.01	13.70
$AGE_{i,t-1}$	-0.406	-1.256	0.000	0.000	1.00	47.81
$TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.28
$NW_{i,t-1}$	-0.000	0.000	0.000	0.000	1.03	0.19
$CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.01	0.50
$K_{i,t-1}$	-1.189	-1.450	-1.191	-0.921	1.00	99.99
$I^N_{i,t-1}$	-0.057	-0.129	-0.067	0.000	1.00	69.05
at the median						
$\Delta Q_{i,t}$	-0.694	-1.328	-0.692	-0.000	1.00	88.41
$V_{i,t}$	-8.543	-12.553	-8.501	-4.842	1.00	99.59
$C_{i,t}$	-9.351	-74.519	-0.000	-0.000	1.01	40.12
$AGE_{i,t-1}$	6.954	-0.000	-0.000	21.451	1.00	48.55
$TL_{i,t-1}$	-0.019	-0.112	-0.000	0.101	1.00	30.87
$NW_{i,t-1}$	-0.005	-0.000	-0.000	-0.000	1.01	1.58
$CL_{i,t-1}$	0.004	-0.041	-0.000	0.085	1.04	17.39
$K_{i,t-1}$	16.778	12.990	16.790	20.445	1.00	99.99
$I^N_{i,t-1}$	0.305	0.065	0.342	0.561	1.00	98.97

Table 5.10. Expected changes in investment

Variable	at mean values	at median values
Change in output ($\Delta Q_{i,t}$)	0.023	0.025
Value of short-term assets ($V_{i,t}$)	0.017	0.017
Cost of capital ($C_{i,t}$)	0.007	0.012
Lagged operator age ($AGE_{i,t-1}$)	-0.004	-0.004
Lagged total liabilities ($TL_{i,t-1}$)	0.000	0.000
Lagged net worth ($NW_{i,t-1}$)	-0.000	0.000
Lagged current liabilities ($CL_{i,t-1}$)	0.000	-0.000
Lagged machinery value ($K_{i,t-1}$)	-0.060	-0.060
Lagged investment ($I^N_{i,t-1}$)	-0.040	-0.063

influence. Lagged operator age, machinery value, and investment have a negative relationship with current investment, where the change in output, the value of short-term assets, and the cost of capital have a positive relationship with investment.

When we compare these figures with the elasticity and expected change estimates from the classical mixed model approach (Table 5.3), we see several important differences. The classical estimates supported stronger relationships between current machinery investment and the cost of capital or lagged operator age. Lagged machinery value had the third or fourth largest effects under the classical estimates, as opposed the largest effects under the Bayesian estimates. Although the signs of the measures are mostly in agreement, the magnitudes of the measures under the two approaches differ significantly, especially the elasticity measures.

5.6 Posterior Predictive Model Checking

One method to evaluate the fit of the model from a Bayesian perspective is to compare the observed data to data generated from the model's posterior predictive distribution. The posterior predictive distribution is defined as

$$p\left(y^{rep}|y\right) = \int p\left(y^{rep}|\theta\right)p\left(\theta|y\right)\partial\theta ,$$

where y^{rep} represents data replicated using the fitted model, y represents the observed data, and θ represents the parameters of the model. It is a posterior distribution since it depends on the observed data and is a predictive distribution since it can generate predictions for possible observations y^{rep}. The basic idea behind posterior predictive analysis is to compare the observed data with replicate data generated by the model. Test quantities, such as test statistics, are defined to measure discrepancies between the model and observed data. A posterior predictive p-value is defined as the probability that the test quantity for the observed data is less than the test quantity for replicated data.

Under the framework used in our analysis, replicate data are easy to produce. We compute the posterior predictive distribution through simulation, employing the last 450 draws from each of the four Gibbs sampler chains from the estimation labeled Main. From each of the draws from the posterior distribution of the parameters, we form a hypothetical investment data set by drawing from

$$I^N_{i,t} \mid y_t, \beta, \mathbf{X}_{i,t}, \sigma_\varepsilon^2, \sigma_y^2, \theta_{i,t} \sim N(y_t + \mathbf{X}_{i,t}\beta, \sigma_\varepsilon^2(\theta_{i,t} + \kappa^2 - \kappa^2\theta_{i,t})).$$

The 1800 replicate data sets summarize the posterior predictive distribution. For the test quantities, we have chosen standard test statistics: the mean, standard deviation, minimum, median, maximum, and the skewness coefficient. The estimated p-values are the proportions of times that the replicate test statistic values exceed the test statistic value obtained from the observed data. Extreme p-values can indicate deficiencies in the model and suggest areas for model improvements.

Figures 5.18-5.23 display the histograms of the test statistic values from the 1800 replicate data sets along with the test statistic value from the observed data. The histograms approximate the posterior predictive distributions for each of the test quantities. The actual test quantity values from the observed data and the posterior predictive p-values are also given in the figures. For the mean and the skewness coefficient, the posterior predictive p-value is zero, indicating that the actual value falls outside of the range of values from the

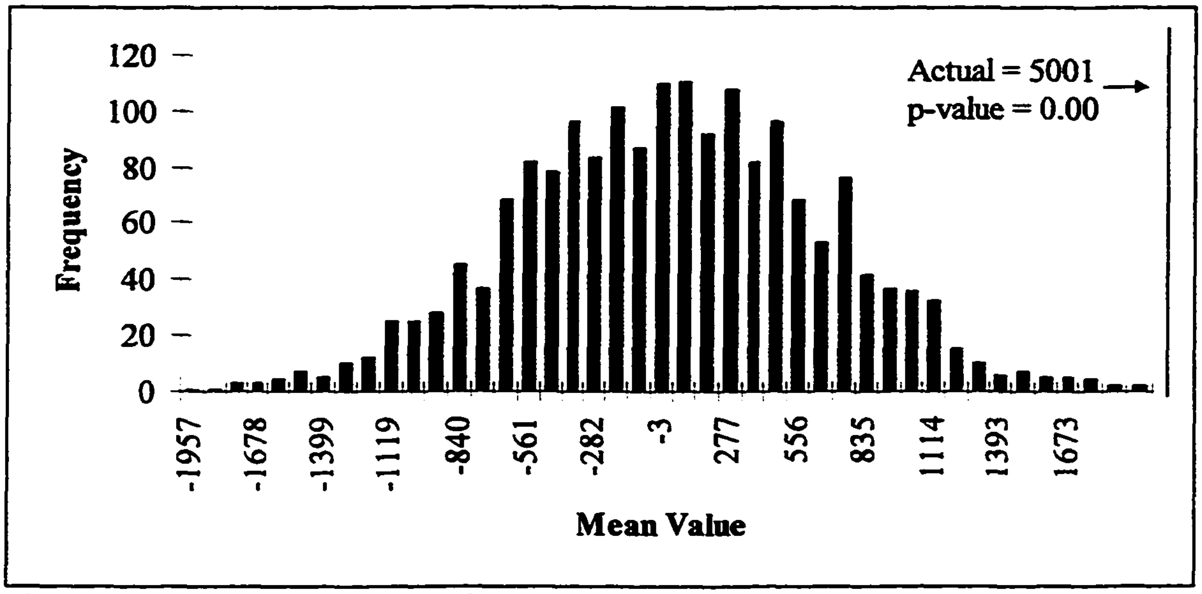


Figure 5.18. Posterior predictive check of the mean

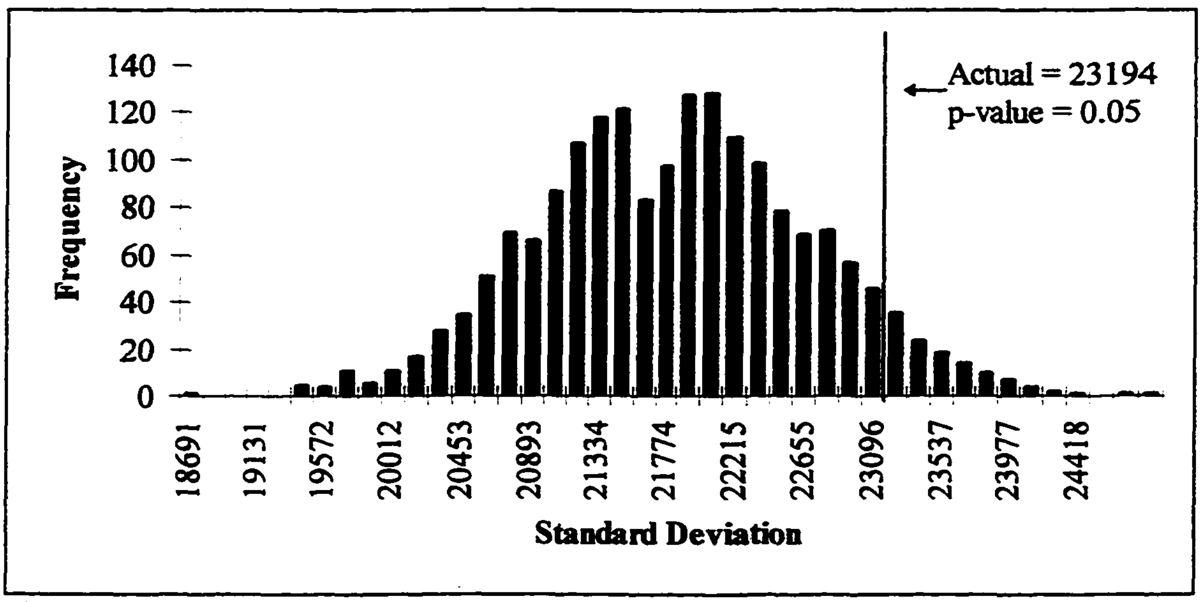


Figure 5.19. Posterior predictive check of the standard deviation

replicate data sets. For the standard deviation and maximum value, the actual value is rarely exceeded by values from the replicate data sets; thus, the posterior predictive p-values are small. For the minimum and median values, the actual value falls well within the range obtained from the replicate data sets.

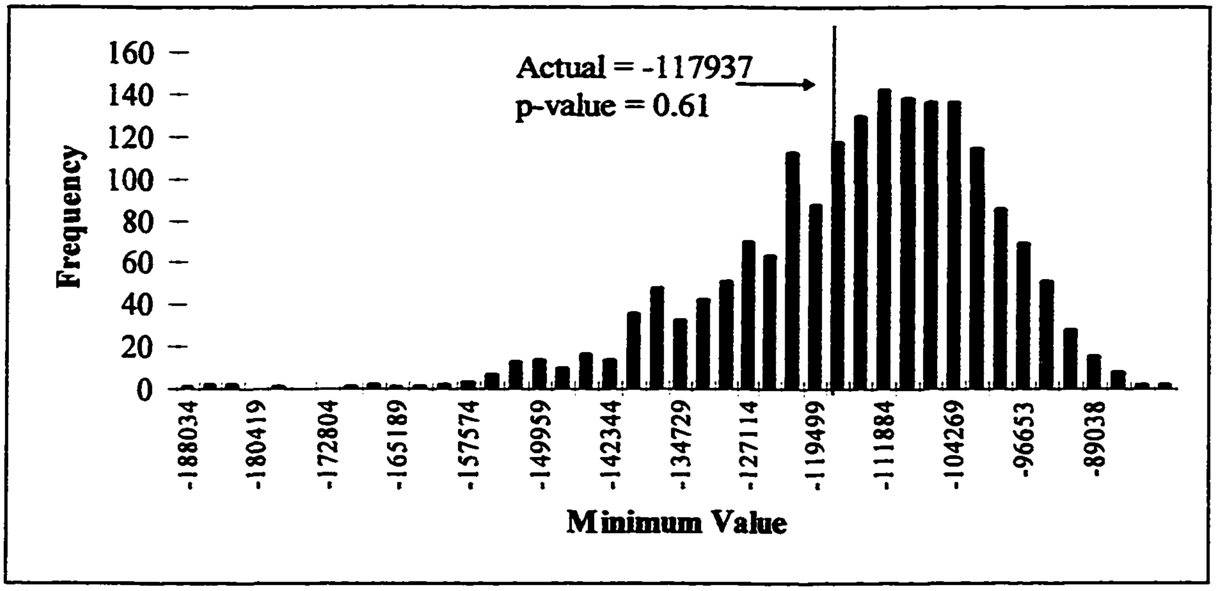


Figure 5.20. Posterior predictive check of the minimum value

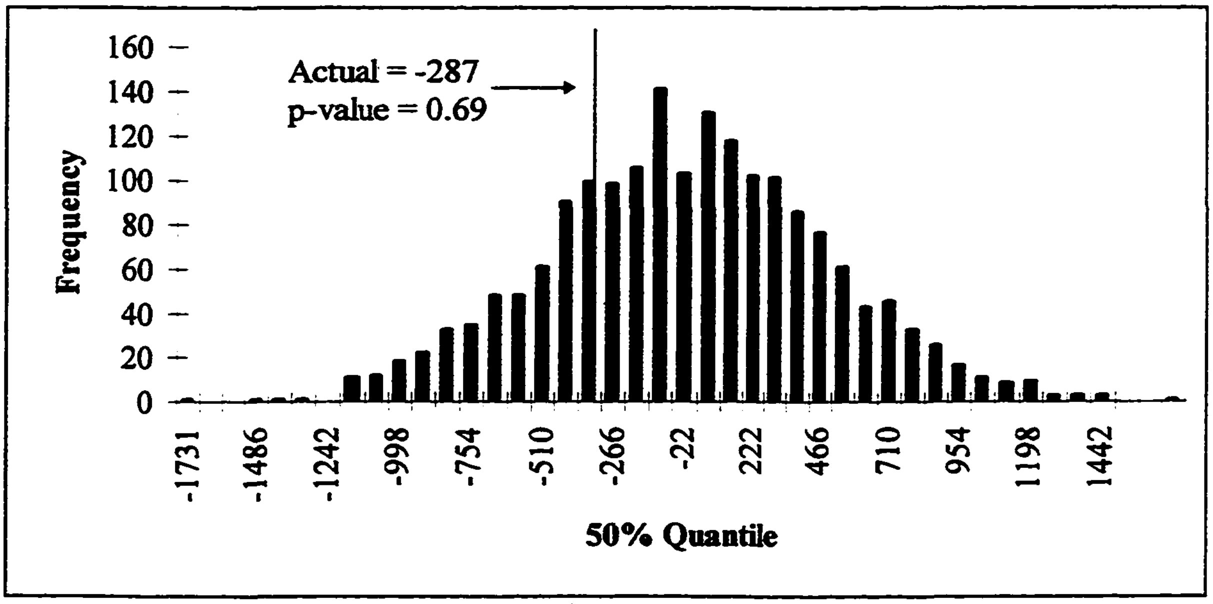


Figure 5.21. Posterior predictive check of the median

These results indicate that the model captures certain aspects (minimum and median values) of the investment data well; but other aspects (mean and maximum values) of the data are inconsistent with the model. The analysis of the skewness coefficient shows that the replicate data sets tended to be symmetric, while the actual data set is skewed. These tests

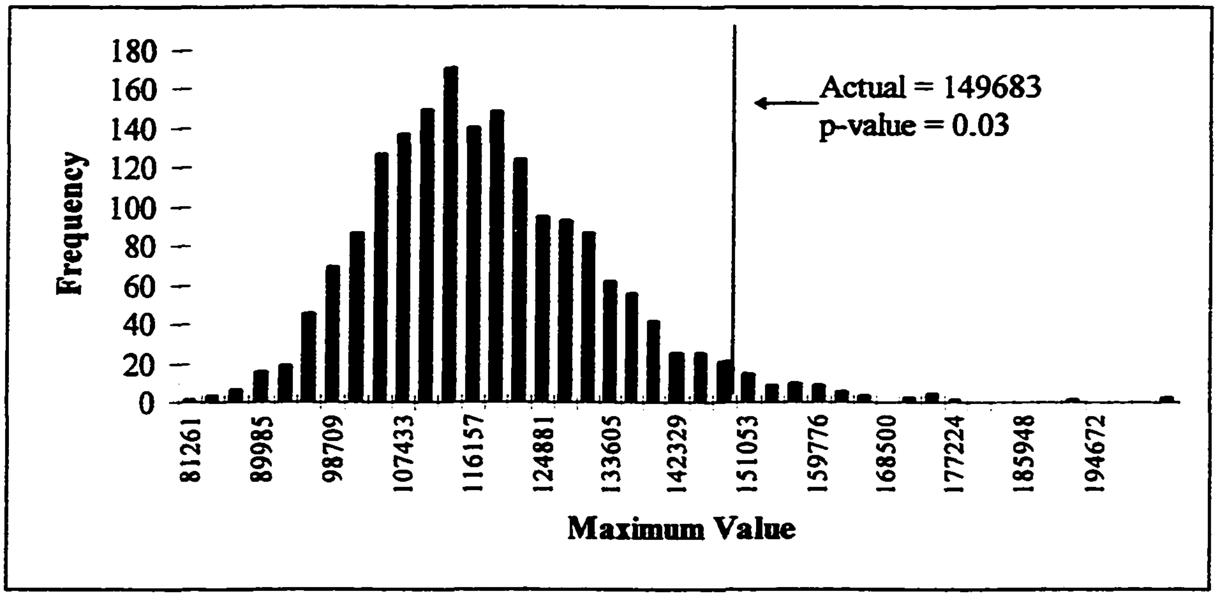


Figure 5.22. Posterior predictive check of the maximum value

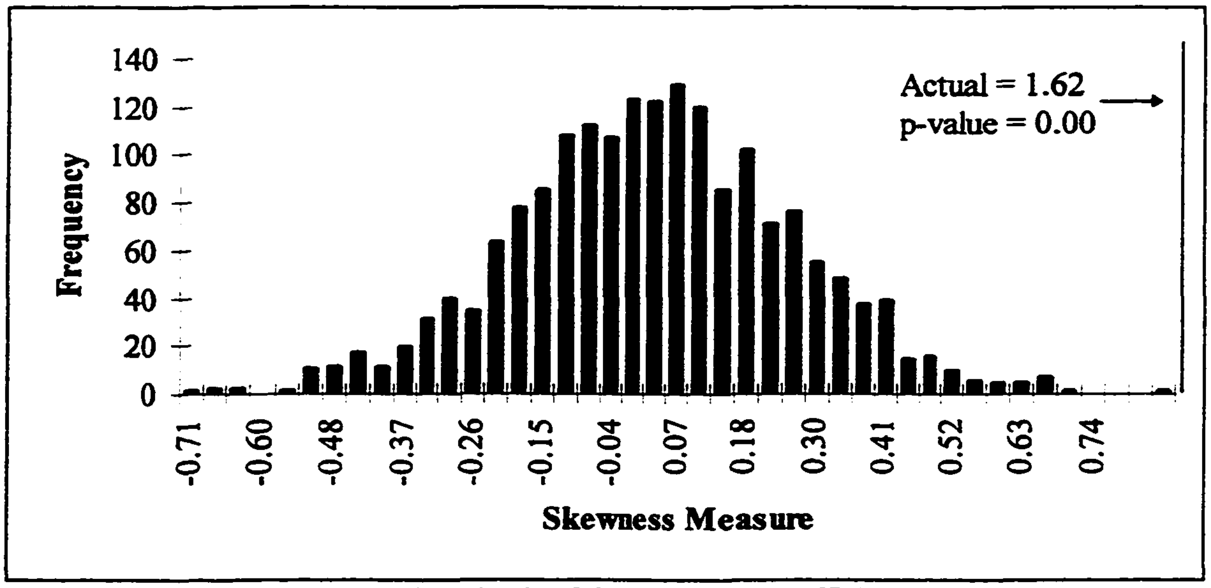


Figure 5.23. Posterior predictive check of the skewness coefficient

indicate that the model might be improved by departing from the normality assumption on the residuals.

One factor that is not included in the investment analysis is farm acreage. To see if farm acreage might add information to the model, we have also examined the correlation between the residuals from the model (based on the 1800 replicate data sets) and the number of acres farmed. Figure 5.24 summarizes the posterior distribution of the correlations. A correlation near zero would imply that farm acreage would add very little new information to the model. However, a non-zero correlation would imply that farm acreage would substantially add to the model. The figure shows very few observations in the vicinity of zero, indicating farm acreage would add information to the model. But the sign of the correlation is indeterminate since the distribution has two distinct pieces. Much of the weight of the posterior distribution is placed in two intervals, (-0.09 to -0.07) and (0.03 to 0.08). Thus, this model check indicates that farm acreage should be included in the next version of the model, but the test does not reveal the direction of the impact farm acreage will have on the model.

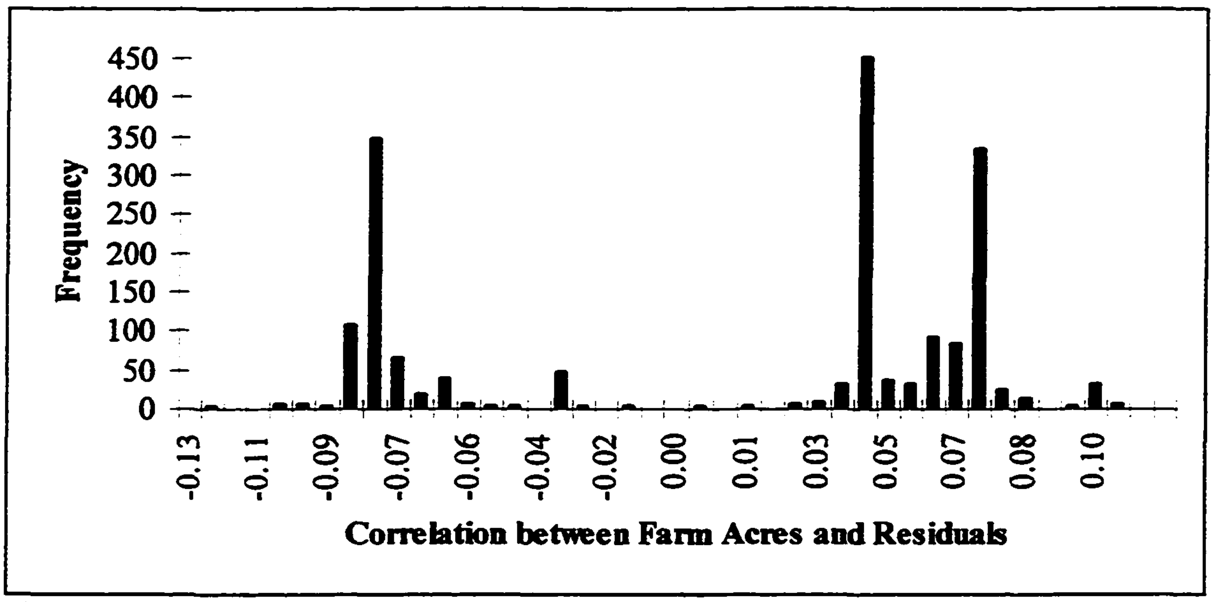


Figure 5.24. Posterior predictive check of the correlation between farm acres and residuals

CHAPTER 6. THE EULER EQUATION APPROACH

6.1 The Euler Equation Approach

The Euler equation approach we take follows Hubbard and Kashyap (1992). To simplify notation, the firm subscript (i) is removed from the following equations. We assume that farmers seek to maximize the present discounted value of investment net cash flows. Assuming all farms face the same prices, farm machinery is the only quasi-fixed input, and machinery is homogeneous, investment net cash flow is defined as

$$\pi_t = p_t F(K_{t-1}, L_{t-1}, N_t) - \mathbf{w}_t{}' \mathbf{N}_t - l_t L_{t-1} - A(I_t, K_{t-1}) - i_{t-1} B_{t-1} + B_t - B_{t-1} - p^I{}_t I_t,$$

where F(.) represents the production function, A(.) represents the adjustment cost function, π is investment net cash flow, p is the output price, K is the capital stock, L represents land, **N** is a vector of variable inputs, **w** is the vector of variable input prices, l is the rental price of land, I represents investment, i is the interest rate, B is total debt, and p^I is the price of investment. Under this specification, it is assumed that this period's machinery investment is not put into use until the next period.[1]

The farmers face several constraints in performing this maximization. Capital accumulation is defined as

$$K_t = (1-\delta)K_{t-1} + I_t ,$$

where δ is the depreciation rate of the capital stock. This period's capital stock value is equal to the sum of the depreciated value of last period's capital stock and this period's

[1] This delay in investment usefulness is often referred to as "time to build". Our time to build is one year.

investment. A transversality condition is assumed to hold so farmers cannot borrow an infinite amount. A positive net cash flow constraint is also included, implying that farms must borrow if cash flow is negative. Here, we present a simplified version of the model by ignoring taxes and inflation, although these factors are taken into consideration when we estimate the Euler equation model.

To arrive at the estimation equation, it is necessary to form, from the objective function and its constraints, the Lagrangian function

$$\mathcal{L} = E_0\left[\sum_{t=1}^{\infty}\left(\prod_{k=0}^{t-1}\beta_k\right)\pi_t + \phi_t\pi_t + \lambda_t\left((1-\delta)K_{t-1} + I_t - K_t\right)\right],$$

where E_t is the expectation operator conditional on information known at time t, β_k is a discount rate, and ϕ_t and λ_t are the Lagrangian multipliers for the non-negative cash flow and capital accumulation constraints, respectively. The Euler equations for the farmer's choice variables are derived. These Euler equations are combined to eliminate the Lagrangian multiplier on the capital accumulation constraint (which represents the shadow value of capital) to arrive at the estimation equation. The Euler equation for investment is given by

$$\frac{\partial \mathcal{L}}{\partial I_t} = (1+\phi_t)\left[-A_{I_t}(I_t, K_{t-1}) - p^I{}_t\right] + \lambda_t = 0.$$

The Euler equations for debt and the capital stock are

$$\frac{\partial \mathcal{L}}{\partial B_t} = (1+\phi_t) - E_t\left[(1+\phi_{t+1})\beta_t(1+i_t)\right] = 0 \text{ and}$$

$$\frac{\partial \mathcal{L}}{\partial K_t} = E_t\left[(1+\phi_{t+1})\beta_t\left\{p_{t+1}F_{K_t}(K_t, L_t, N_{t+1}) - A_{K_t}(I_{t+1}, K_t)\right\} + \beta_t(1-\delta)\lambda_{t+1}\right] - \lambda_t = 0,$$

respectively. The Euler equations for investment and debt are substituted into the capital stock Euler equation, to create the estimation equation

$$p_{t+1}F_{K_t}\left(K_t,L_t,N_{t+1}\right)-A_{K_t}\left(I_{t+1},K_t\right)+(1-\delta)\left[A_{I_{t+1}}\left(I_{t+1},K_t\right)+p^I{}_{t+1}\right]$$
$$-\left(1+i_t\right)\left[A_{I_t}\left(I_t,K_{t-1}\right)+p^I{}_t\right]=\eta_{t+1}$$

where η_{t-1} is the expectation error. The expectations are evaluated at realized values and an expectation error is added to the equation. Under the rational expectations assumption, this expectation error has a mean of zero ($E_t[\eta_{t+1}] = 0$) and is uncorrelated with any information to which the farmer has access at the time of making the decision.

The first three terms in the equation above represent the marginal benefit of investing this period. The first term is the (next period's) marginal revenue from making the investment in this period. The second term is the additional adjustment costs from investing in this period. The third term is the marginal cost of investment in the next period. The last two terms in the equation represent the marginal cost of investing during this period. The fourth term is the marginal adjustment cost from an additional dollar of investment. The fifth term is the price paid per unit of capital bought this period.

Estimation is performed to obtain values for the parameters in the production and adjustment cost functions. For the production function, most studies of this kind have followed the convention of the Q literature and have assumed that average and marginal products of capital are equal. Marginal and average products of capital are equal when the production function is homogeneous of degree one in machinery and variable inputs, no quantity constraints exist in either the input or output markets, and farmers are price takers.

This also implies perfect competition and constant returns to scale. The typical assumption for the marginal value product of capital is given by

$$p_{t+1}F_{K_t} = \frac{Y_{t+1} - C_{t+1}}{K_t},$$

where Y_{t+1} is gross revenue from production and C_{t+1} is variable cost.

Specifications for the adjustment cost function have also followed the Q literature. Adjustment costs are taken to be linearly homogeneous in investment and capital, so as to equate average and marginal Q. The Q model uses a quadratic cost function where costs are incurred when investment deviates from a given rate. Often the adjustment cost function is given as

$$A(I_t, K_{t-1}) = 0.5\theta_0\left(\frac{I_t}{K_{t-1}} - \upsilon\right)^2 K_{t-1},$$

where υ represents an average or "normal" investment rate and θ_0 is an unknown parameter to be estimated.

To allow for asymmetries in adjustment costs, we have specified the adjustment cost function as a piecewise quadratic function with the breakpoint at $\frac{I_t}{K_{t-1}} = \upsilon$ with θ_1 representing adjustment costs when $\frac{I_t}{K_{t-1}} < \upsilon$ and θ_2 representing adjustment costs when $\frac{I_t}{K_{t-1}} \geq \upsilon$. The basic model we employ in the Euler analysis is the estimation equation with the marginal value product of capital as depicted above and a standard symmetric adjustment cost function. We also estimate the parameters in the model using the asymmetric adjustment cost specification described above.

6.2 Incorporation of Financial Constraints in the Euler Approach

Several studies have included a financial constraint to the Euler equation investment model. The financial constraint is of the form $B^*_t \geq B_t$; the farmer is constrained to have outstanding debt, B_t, less than or equal to some debt ceiling, B^*_t. Hubbard and Kashyap (1992) suggest B^*_t could be made a function of net worth. The financial constraint we employ incorporates the "5 Cs" of lending: character, capacity (cash flow), collateral, credit rating, and capital (owner's equity). The expanded Lagrangian function with the debt constraint is given below:

$$\mathcal{L} = E_0\left[\sum_{t=1}^{\infty}\left(\prod_{k=0}^{t-1}\beta_k\right)\pi_t + \phi_t\pi_t + \lambda_t\left((1-\delta)K_{t-1} + I_t - K_t\right) + \omega_t\left(B^*_t - B_t\right)\right],$$

where ω_t is the Lagrangian multiplier for the debt constraint. Following the same procedure as before, we arrive at the estimating equation for this variation of the model:

$$\begin{aligned}(1-\hat{\omega}_t)\left\{p_{t+1}F_{K_t}(K_t, L_t, N_{t+1}) - A_{K_t}(I_{t+1}, K_t) + (1-\delta)\left[A_{I_{t+1}}(I_{t+1}, K_t) + p^I{}_{t+1}\right]\right\} \\ -(1+i_t)\left[A_{I_t}(I_t, K_{t-1}) + p^I{}_t\right] = \eta_{t+1},\end{aligned}$$

where $\hat{\omega}_t = \dfrac{\omega_t}{1+\phi_t}$.

The parameter $\hat{\omega}_t$ has been modeled in a variety of ways. Hubbard and Kashyap (1992) assume it is a multiple of the change in net worth. Bierlen (1994) takes it as a quadratic function of the level of net worth. Whited (1992) specifies $\hat{\omega}_t$ as a quadratic function of the firm's debt to asset ratio and the ratio of the firm's interest expenses to the sum of the interest expenses and cash flow. Hubbard, Kashyap, and Whited (1995) model $\hat{\omega}_t$ as a function of the firm's cash flow and an interest rate spread.

Due to its structure, $\hat{\omega}_t$ must be non-negative. The specifications above may not reflect this point. To capture this restriction, we form $\hat{\omega}_t$ as an exponential function where the power of the exponential will depend on the 5 Cs of lending. This model specification adds to the existing literature by expanding the financial constraint specification.

Another approach to incorporate a financial constraint into the model is to model the interest rate the farmer faces. In a paper examining scale economies in banks, Hughes and Mester (1995) model loan interest rates as the product of a risk-free interest rate and a risk premium. The risk premium is a function of the bank's outputs, capitalization, and risk structure. This same approach can be taken from the farmer's point of view. The interest rate the farmer faces is composed of a risk-free interest rate and a risk premium. This approach (also referred to as an elastic credit supply approach) has been employed by Bond and Meghir (1994), Estrada and Vallés (1995), and Barran and Peeters (1998). This risk premium is likely to depend on the 5 Cs of lending. Higher risk premiums would translate into higher loan interest rates which would effectively prohibit borrowing, thus constraining the farmer's investment choices.

The financial constraint would then take the form of an interest rate constraint, such as $i_t = i^*_t\mu_t$ where i_t is the loan interest rate, i^*_t is the risk-free interest rate, and μ_t is the risk premium ($\mu_t \geq 1$). This constraint can be directly inserted into the Lagrangian function. In order to meet the range restriction on μ_t, we model μ_t as one plus an exponential where the power of the exponential will depend upon the 5 Cs of lending.

Both of these approaches to the financial constraint can be brought into one specification, but previous attempts to do so have not succeeded. Convergence problems have been encountered in employing the interest rate specification (Barran and Peeters 1998). Due to this difficulty and the limited time frame of the data set, we have chosen to focus on the debt ceiling approach.

We have chosen to model $\hat{\omega}_t$ as an exponential function of the 5 Cs. Here,

$$\hat{\omega}_t = \exp[\rho_0 + \rho_1 AGE_{t-1} + \rho_2 V_t + \rho_3 TL_{t-1} + \rho_4 NW_{t-1} + \rho_5 CL_{t-1} + \rho_6 (AGE_{t-1})^2 + \rho_7 (V_t)^2 + \rho_8 (TL_{t-1})^2 + \rho_9 (NW_{t-1})^2 + \rho_{10} (CL_{t-1})^2],$$

where AGE is the operator's age, V is the value of short-term assets, TL is the total farm liability, NW is the net worth, CL is the short-term farm liability, and the ρs are unknown parameters to be estimated from the data. Character is proxied by the farmer's age. Capacity is proxied by the value of short-term assets (feeder livestock, stored crops, supplies, fertilizer, and cash). Net worth, total liabilities, and short-term liabilities (due within one year) complete the equation. Both linear and quadratic terms are included in the specification. Age, net worth, and both liability variables are lagged to represent the financial state of the farm before the investment was undertaken. The concurrent value of short-term assets is used to represent the farm's potential cash flow. To control for farm size, all of the financial variables in the debt ceiling constraint are divided by the lagged value of the capital stock.

We have added fixed firm and time effects to the estimating equation. The firm effects capture other firm characteristics not included in the model. The time effects capture aggregate business cycle components common to all agents. The complete estimation equation is given by

$$\left(1-\hat{\omega}_{i,t}\right)\left\{p_{t+1}F_{K_t}\left(K_{i,t},L_{i,t},N_{i,t+1}\right)-A_{K_t}\left(I_{i,t+1},K_{i,t}\right)+(1-\delta)\left[A_{I_{t-1}}\left(I_{i,t+1},K_{i,t}\right)+p^{I}_{i,t+1}\right]\right\}$$
$$-\left(1+i_{i,t}\right)\left[A_{I_t}\left(I_{i,t},K_{i,t-1}\right)+p^{I}_{i,t}\right]+f_i+y_t=\eta_{i,t+1},$$

where the f_is are the firm effects and the y_ts are the time effects. To account for the effects of taxes and inflation, we modify the interest rate and the price of investment. The interest rate is computed as $i_t = (1-\tau_{i,t})i^{*}_{t} - \psi_t$ where $\tau_{i,t}$ is the marginal federal tax rate, i^{*}_{t} is the average effective interest rate on all non real estate agricultural loans reported by the Federal Reserve, and ψ_t is the percent change in the Gross Domestic Product deflator reported by the Bureau of Economic Analysis. The price of investment, p^{I}_{t}, is computed as $(1-\tau_{i,t}z_{i,t})p^{I^{*}}_{t}$ where $z_{i,t}$ is the present value of future depreciation deductions from investment at time t and $p^{I^{*}}_{t}$ is the price index for farm machinery reported by the United States Department of Agriculture. The variable $z_{i,t}$ is calculated as $\dfrac{\delta}{\delta+\dfrac{i^{*}_{t}-\psi_t}{1-\xi}}$, where ξ is the accrual equivalent tax rate on capital gains. Following Whited (1992) and Bierlen (1994), ξ is set equal to 0.05.

6.3 Estimation Technique

The estimation technique employed in most of the Euler equation investment literature is an instrumental variable generalized method of moments (IV-GMM) approach. There are several reasons why this technique has been chosen. First, the IV-GMM technique incorporates the rational expectations assumption directly. Second, the technique yields consistent parameter estimates if the instruments are uncorrelated with the stochastic error. Third, the parameter estimates can be produced to be robust to heteroscedasticity and serial correlation. Fourth, there are simultaneity problems within the model; IV-GMM can handle such problems.

In method of moments estimation, sample moments are equated with population moments and a solution is obtained for the population parameters. For example, given a sample of independent draws from a distribution, we can equate any moment of the distribution to the corresponding sample moment. For the population mean, the method of moments estimator is the sample mean. Generalized method of moments (GMM) estimation extends the ordinary method of moments technique by utilizing both conditional and unconditional moments and by possibly having these moments depend on unknown parameters (Davidson and MacKinnon, 1993).

The just identified[2] IV-GMM estimator is described below. Suppose we are given $\mathbf{y}_t$, an n x 1 vector of dependent variables, and **X**, an n x k matrix of independent variables, and

[2] In the just identified case, the numbers of instruments is equal to the number of parameters. In the overidentified case, the number of instruments exceeds the number of parameters.

are attempting to estimate β, a k x 1 unknown parameter vector. Let $\mathbf{u}_t = f_t(\mathbf{y}_t, \mathbf{X}, \beta)$ be a function with an expectation of zero. The equations which define the GMM estimator for this problem are given by $\frac{1}{n}\Sigma_{t=1}^{n} f_t(\mathbf{y}_t, \mathbf{X}, \beta) = \frac{1}{n}\Sigma_{t=1}^{n} \mathbf{u}_t = 0$. The GMM estimator of β, $\hat{\beta}$, mimics the moment restrictions by setting $\hat{\beta}$ to minimize the following quadratic form,

$J_n(\beta) = \left(\frac{1}{n}\Sigma_{t=1}^{n}\mathbf{u}_t\right)' \mathbf{W}_n\left(\frac{1}{n}\Sigma_{t=1}^{n}\mathbf{u}_t\right)$, where $\mathbf{W}_n$ is a positive semidefinite matrix with $\lim_{n\to\infty} \mathbf{W}_n = \mathbf{W}$, a positive definite matrix (Ogaki, 1993). If $\mathbf{u}_t$ is serially uncorrelated, then the optimal GMM estimator is reached with $\mathbf{W} = \Omega^{-1} = (E[\mathbf{u}_t\mathbf{u}_t'])^{-1}$ (Hansen, 1982). In our case, the function is the product of the instruments, **Z**, and the expectational error, η_t. Under rational expectations, the errors in the expectations of the economic agents should be independent of all variables within their information sets, $E[\mathbf{Z}\eta_t] = 0$. Thus, the product of the instruments and errors can serve as both the embodiment of the rational expectations assumption and the moment conditions needed to employ the IV-GMM approach.

Hansen (1982) examines the large sample properties of GMM estimators. He shows that GMM estimators are strongly consistent and asymptotically normal given that the observable variables are stationary and ergodic. Also, many econometric estimators, such as ordinary least squares and instrumental variables approach, can be represented as special cases of GMM estimators (Davidson and MacKinnon, 1993). Fixed firm and time effects can be dealt with in an efficient manner in GMM estimation (Lahiri, 1993). Hansen and Singleton (1982) represents the first application of IV-GMM to a rational expectations, Euler equation model. Since then, IV-GMM has become the technique of choice for these types of models.

If the number of instruments used (q) in the IV-GMM approach exceeds the number of parameters (k), then the system is labeled as overidentified. This indicates that the number of moment conditions used to estimate the system exceeds what is required. A chi-square test statistic can be employed to test these overidentifying conditions, using a test often referred to as the Sargan or Hansen's J-test. In his 1982 paper, Hansen showed that the product of the number of observations and the minimized value of the GMM objective function, $nJ_n(\hat{\beta})$, has an asymptotic χ^2 distribution with q–k degrees of freedom when $\mathbf{W} = \Omega^{-1}$. When we examine variations of the model (i.e., look at parameter restrictions), we test these using a likelihood ratio type of test statistic, $n(J_n(\hat{\beta}_r) - J_n(\hat{\beta}))$ where $\hat{\beta}_r$ is the GMM estimator for the restricted version of the model and $\hat{\beta}$ is the GMM estimator for the unrestricted version. Under a set of regularity conditions and the use of the same estimator for Ω in both the restricted and unrestricted versions of the model, this test statistic has an asymptotic χ^2 distribution with s degrees of freedom where s is the number of restrictions.

The estimation is performed using the GMM procedure in TSP 4.2B. The data set is the same as was used for the composite regression approach and is described in Section 4.1. Following White (1980), the computed standard errors are consistent when disturbances are heteroscedastic. Annual equations with cross-equation restrictions are estimated. Due to the presence of lagged dependent variables and fixed firm effects, the equations are first-differenced to remove the fixed firm effects and all instruments are lagged one period. The instruments are operator age, the price of investment, the marginal value product of capital, farm total liability, net worth, farm short-term liability, short-term assets, non-farm income, tax expenses, and interest expenses. All financial instruments are divided by the value of the capital stock to control for farm size. The parameter υ is set at 0.135, the average investment rate over the entire sample. The depreciation rate, δ, is set at 0.10, the same figure the Iowa Farm Business Association employs. Since we include squared terms in the financial constraint, we have removed the means of these variables to alleviate possible multicollinearity.

During the estimations, we begin by estimating the parameters of the model under four different specifications. We have labeled these specifications Models 1-4. Model 1 is the basic Euler investment equation with symmetric adjustment costs. Model 2 extends this basic model to have asymmetric adjustment costs. Model 3 adds the financial constraint to the basic model. Model 4 adds the constraint to Model 2. We tested for farm effects and found them to be significant; thus, we proceed to estimate the parameters of the models including these effects.

CHAPTER 7. EULER EQUATION RESULTS, EXTENSIONS, AND DISCUSSION

7.1 Euler Equation Results

The results for the investment equations in first-difference form are presented in Table 7.1. The four specifications cover the interactions between symmetric and asymmetric adjustment costs and the inclusion (exclusion) of the financial constraint. For statistical significance, we use the five percent level for both the parameter estimates and the model tests. The J-statistic tests labeled (O. R.) are the standard tests (often referred to as Sargan or Hansen's J-tests) for GMM estimation. A p-value above 0.05 indicates the model is not rejected. A p-value below 0.05 indicates evidence for rejecting the model. The tests labeled (vs. M 1) and (vs. M 4) are comparing nested models. A p-value above 0.05 indicates the more restrictive model is not rejected in favor of the less restrictive model. A p-value below 0.05 indicates evidence for rejecting the more restrictive model in favor of the less restrictive model.

In all four cases the model is rejected. But among these models, the restricted specification with symmetric adjustment costs and no financial constraint (Model 1) is the preferred model. In each case, the more sophisticated model is rejected for its more restrictive counterpart. In three cases, adjustment costs are estimated to be negative, and two of these are significantly different from zero. The addition of the financial constraint to the model has an unique effect on the adjustment cost parameters. When the constraint is added to the model, the adjustment cost parameter estimates are reduced by nearly a factor of 100 and they reverse signs. In comparing these results to others in the field, we find

Table 7.1. Parameter estimates for investment equations in first-difference form

	Model 1: sym. adj. costs and no fin. constraint		Model 2: asym. adj. costs and no fin. constraint		Model 3: sym. adj. costs with fin. constraint		Model 4: asym. adj. costs with fin. constraint	
Parameter	Value	S. E.	Value	S. E.	Value	S. E.	Value	S. E.
Adj. costs								
θ_0 (Sym.)	-0.8766^*	0.0045			0.0078	0.0043		
θ_1 (Asym.)			1.32^*	0.14			-0.044	0.060
θ_2 (Asym.)			-0.9960^*	0.0079			0.0106^*	0.0043
Fin. constraint								
ρ_0					-0.002	0.011	-0.004	0.011
ρ_1 (AGE_{t-1})					-0.00047	0.00049	-0.00041	0.00049
ρ_2 (V_t)					0.0010^*	0.00051	0.00115^*	0.00055
ρ_3 (TL_{t-1})					-0.00038	0.0012	-0.0005	0.0013
ρ_4 (NW_{t-1})					-0.00073^*	0.00018	-0.00079^*	0.00018
ρ_5 (CL_{t-1})					0.0042	0.0027	0.0046	0.0026
ρ_6 $(AGE_{t-1})^2$					2.0×10^{-5}	2.5×10^{-5}	2.4×10^{-5}	2.5×10^{-5}
ρ_7 $(V_t)^2$					0.6×10^{-5}	1.0×10^{-5}	0.4×10^{-5}	1.1×10^{-5}
ρ_8 $(TL_{t-1})^2$					-2.1×10^{-5}	2.6×10^{-5}	-1.6×10^{-5}	2.8×10^{-5}
ρ_9 $(NW_{t-1})^2$					$2.54\times10^{-6*}$	7.9×10^{-7}	$2.71\times10^{-6*}$	7.8×10^{-7}
ρ_{10} $(CL_{t-1})^2$					-1.58×10^{-4}	9.6×10^{-5}	-1.70×10^{-4}	9.4×10^{-5}
J-stat. tests[a]		p-value		p-value		p-value		p-value
χ^2 (O. R.)	86.19	0.0027	83.53	0.0036	73.84	0.0017	73.59	0.0013
d. f.	53		52		42		41	
χ^2 (vs. M 1)			2.67	0.1023	12.35	0.3376	12.61	0.3984
d. f.			1		11		12	
χ^2 (vs. M 4)			9.94	0.5361	0.25	0.6163		
d. f.			11		1			

[a]O. R. stands for overidentifying restrictions, M 1 stands for Model 1, and M 4 stands for Model 4. The parameter estimates for the time effects are not presented.
*Significantly different from zero at the 5% level based on the two-tailed t-statistic.

some similarities, but many differences. For comparison purposes, we concentrate on the Euler equation models with the financial constraint since all of the other studies found results favorable to that specification. Based on the mean values of $\left(\frac{I_{i,t}}{K_{i,t-1}} - \upsilon\right)^2$ from the data set and the adjustment cost parameter estimates from Models 3 and 4 above, adjustment costs are quite small, –0.19 percent to 1.08 percent of the value of the pre-existing capital stock. Hubbard and Kashyap (1992), Whited (1992), and Bierlen (1994) found much higher adjustment costs, between 10 and 15 percent. Barren and Peeters (1998) and Estrada and Vallés (1995) have adjustment costs similar to what is shown here. In their paper, Estrada and Vallés suggest a possible reason for this discrepancy is measurement error due to the fact that investment is constructed from capital stock changes. Our investment series was calculated using this technique; thus, this may explain the adjustment cost results. We examine this issue in the next section.

Several of the financial constraint variables have significant coefficients. The multiplier for the financial constraint can be thought of as the shadow value of borrowing or external finance, the value of an additional unit of debt. Given mean values for all of the terms in the financial constraint, the shadow value of external finance is 100 percent. Bierlen found a mean value of 69 percent for the shadow value of external finance on a similar agricultural panel data set. Using a manufacturing panel data set, Whited computed a median value of 12 percent. In both of these studies, there were firms that had computed shadow values of external finance near 100 percent. Chirinko (1993) points out that in a number of cases, studies have shown shadow values of external finance greater than 100 percent. Our study also indicates the possibility of shadow values above 100 percent. Such high shadow values imply that farms face significant financing constraints and possible credit rationing.

Net worth and the value of short-term assets were the two variables that were significant in the financial constraint. To explore the relationships implied between the shadow value of finance and these variables, we have graphed the shadow value over relevant ranges of values for these variables. These graphs are given in Figures 7.1 and 7.2. The graphs are based on the parameter estimates from Model 3 with all other variables set at their mean values. Both graphs are basically linear. The quadratic effect for net worth, although statistically significant, is so small as to have a negligible impact. As the ratio of the values of short-term assets to the capital stock increases, the shadow value of external finance also increases. This result differed from what was expected. Most studies, such as Hubbard, Kashyap, and Whited (1995), have found the relationship between variables representing the capacity of the firm to carry financing and the shadow value of external

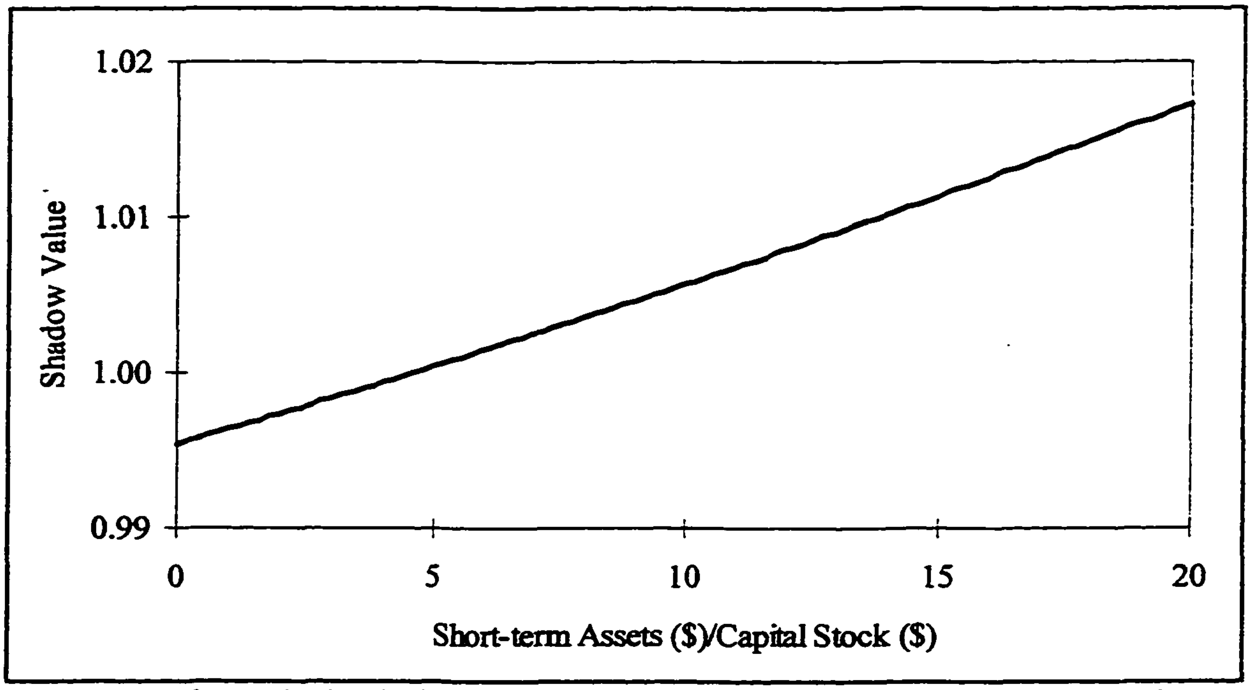


Figure 7.1. Change in the shadow value of external finance due to short-term asset value

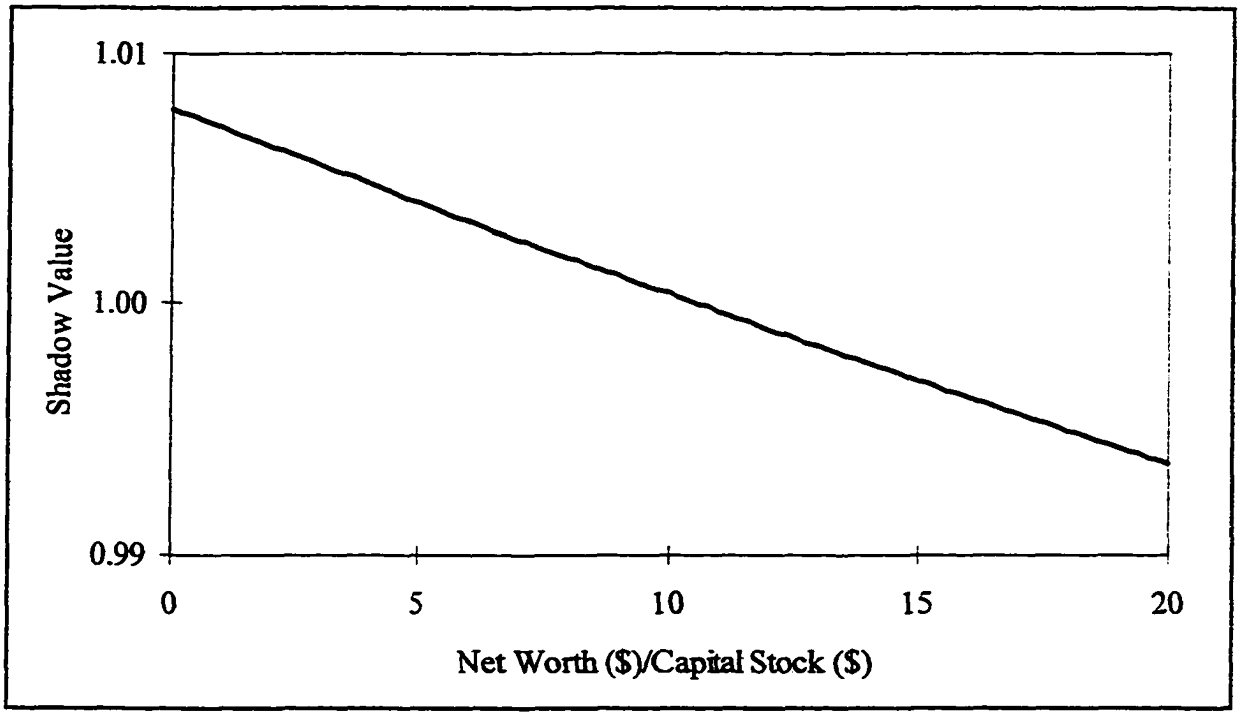


Figure 7.2. Change in the shadow value of external finance due to net worth

finance to be inverse. Firms that sustain a higher cash flow are less likely to fail in their debt obligations. However, our results indicate farms with higher short-term asset values are more constrained in the debt market. One argument for this may be a signaling argument. Lenders may interpret higher short-term asset values as a signal that the farmer is unwilling to liquidate their own assets to invest in the project, indicating a more risky venture. Another argument is that the farmer may be practicing internal credit rationing, foregoing possible loans and building up farm reserves for investment in the future.

Net worth and the shadow value of external finance have an inverse relationship. As net worth increases, the shadow value of external finance decreases. The significance of the quadratic term indicates that the shadow value decreases at a decreasing rate. Both Hubbard and Kashyap (1992) and Bierlen (1994) found net worth to be significant. For comparison purposes, the Bierlen study is the most closely related. In his study, Bierlen found similar results for this relationship.

7.2 Examining Gross Investment

In the discussion of the adjustment cost results from the previous section, we mentioned that a possible explanation may be measurement error due to investment being calculated from changes in the value of the capital stock. Many investment studies have been concerned with this issue and have chosen to compute the series for the value of capital. Following a technique outlined by Salinger and Summers (1983), given an initial value of the capital stock, the gross investment (capital purchases) series, the price of investment, and the estimated life left in the capital stock, the series for the value of the capital stock is computed. The estimated life of the capital stock, L_t, in any year is equal to $L_t = \frac{K_{t-1} + I_t}{DE_t}$ where K_{t-1} is the previous year's value of capital, I_t is the current year's gross investment, and DE_t is the current year's capital depreciation. The average useful life, L, is often employed in the formulation. Then given the average capital life, L, an initial value of the capital stock, K_0, and the gross investment series, I_t, the capital stock value series is computed as $K_t = \left(K_{t-1}\frac{p^I{}_t}{p^I{}_{t-1}} + I_t\right)\left(1 - \frac{2}{L}\right)$ where p^I is the price of the capital good. The assumptions underlying this technique are:

1) all of the capital stock has the same amount of useful life (L);
2) book depreciation is calculated using the straight-line method;
3) actual depreciation is exponential with a depreciation rate of 2/L; and

4) all investment is made at the beginning of the year and all depreciation is taken at the end of the year.

The investment series that is examined in most studies is gross investment, not net investment. We had chosen to examine net investment due to the length and composition of our data set. Since the data set only contained five years of data, losing one year to have lead or lagged variables is a high cost. Also, gross investment and depreciation were only reported in three of those years. The examination of net investment would leave us with three years of data to explore, while an examination of gross investment would limit us to two years of data. However, given the results of the previous section, we now examine gross investment under the Euler equation framework, calculating the value of the capital stock using the Salinger and Summers method.

The structure of the Euler equation model remains the same. Gross investment is equal to machinery and equipment purchases. Depreciation is listed as economic depreciation for machinery and equipment. We allow L, the useful life of the capital stock, to vary by farm. The results for gross investment Euler equations are given in Table 7.2. Firm effects are included in the models; thus, the equations are first-differenced during the estimation.

Three of the four models are not rejected; only Model 1, the most restrictive case with symmetric adjustment costs and no financial constraint, is rejected. This indicates that the overall fit has improved by examining gross investment, instead of net investment. Among the four models, the model with symmetric adjustment costs and a financial constraint, Model 3, is the preferred model. Many of the patterns seen in the net investment

Table 7.2. Parameter estimates for gross investment equations

	Model 1: sym. adj. costs and no fin. constraint		Model 2: asym. adj. costs and no fin. constraint		Model 3: sym. adj. costs with fin. constraint		Model 4: asym. adj. costs with fin. constraint	
Parameter	Value	S. E.	Value	S. E.	Value	S. E.	Value	S. E.
Adj. costs								
θ_0 (Sym.)	-0.32	0.26			0.063	0.068		
θ_1 (Asym.)			1.8	2.0			-0.22	0.59
θ_2 (Asym.)			-0.62	0.43			0.10	0.11
Fin. constraint								
ρ_0					0.016	0.039	0.015	0.039
ρ_1 (AGE_{t-1})					-0.0023	0.0030	-0.0020	0.0029
ρ_2 (V_t)					0.0086*	0.0043	0.0073	0.0051
ρ_3 (TL_{t-1})					0.0017	0.0032	0.0015	0.0032
ρ_4 (NW_{t-1})					-0.0035	0.0022	-0.0025	0.0030
ρ_5 (CL_{t-1})					0.0072	0.0090	0.0058	0.0099
ρ_6 $(AGE_{t-1})^2$					$-1.9\text{x}10^{-4}$	$1.7\text{x}10^{-4}$	$-1.6\text{x}10^{-4}$	$1.7\text{x}10^{-4}$
ρ_7 $(V_t)^2$					$-2.8\text{x}10^{-4*}$	$1.3\text{x}10^{-4}$	$-2.4\text{x}10^{-4}$	$1.5\text{x}10^{-4}$
ρ_8 $(TL_{t-1})^2$					$7.4\text{x}10^{-5}$	$7.7\text{x}10^{-5}$	$6.3\text{x}10^{-5}$	$8.1\text{x}10^{-5}$
ρ_9 $(NW_{t-1})^2$					$9.4\text{x}10^{-5*}$	$4.1\text{x}10^{-5}$	$7.9\text{x}10^{-5}$	$5.3\text{x}10^{-5}$
ρ_{10} $(CL_{t-1})^2$					$-1.8\text{x}10^{-3}$	$9.3\text{x}10^{-4}$	-0.0016	0.0011

J-stat. tests[a]		p-value		p-value		p-value		p-value
χ^2 (O. R.)	31.10	0.0194	24.42	0.0807	4.72	0.5796	4.68	0.4559
d. f.	17		16		6		5	
χ^2 (vs. M 1)			6.68	0.0098	26.37	0.0057	26.42	0.0094
d. f.			1		11		12	
χ^2 (vs. M 4)			19.74	0.0490	0.04	0.8358		
d. f.			11		1			

[a]O. R. stands for overidentifying restrictions, M 1 stands for Model 1, and M 4 stands for Model 4. The parameter estimates for the time effects are not presented.
[*]Significantly different from zero at the 5% level based on the two-tailed *t*-statistic.

estimates also occur in the gross investment estimates. In three of the models, we have negative adjustment costs, although none of the adjustment cost parameter estimates are statistically significant. The addition of the financial constraint leads to a reduction and sign switch for the adjustment cost estimates. This time the reduction is by a factor of between five and nine times. Based on the mean values of $\left(\frac{I_{i,t}}{K_{i,t-1}} - \upsilon\right)^2$ from the data set and the adjustment cost parameter estimates from Models 3 and 4 above, adjustment costs are still quite small, –0.42 percent to 2.58 percent of the value of the capital stock.

Both net worth and the value of short-term assets are found to be statistically significant in the preferred model, Model 3. Given mean values for all of the terms in the financial constraint, the shadow value of external finance is 102 percent, again implying that farms face financial constraints and possible credit rationing. If we compare the financial constraint parameter estimates from Model 3 for both the gross and net investment runs, we find that the same two variables appear significant, although the pattern has changed. In the net investment estimation, both the linear and quadratic term for net worth and the linear term for the value of short-term assets are significant. In the gross investment estimation, both terms for the value of short-term assets and the quadratic term for net worth are significant. Signs change for the parameter estimates of the quadratic terms for operator age and the value of short-term assets and both terms of total liabilities. The parameter estimates for the financial constraint from the gross investment estimation also are more palatable from an economic interpretation viewpoint. For example, under the net investment parameter estimates, total liabilities and shadow value of external finance have an inverse relationship, indicating that credit restrictions ease with higher levels of debt. This is counter to what we had expected. The gross investment parameter estimates show total liabilities and the shadow value of external finance to have a direct relationship, higher debt levels are paired with tighter credit restrictions.

In Figures 7.3 and 7.4, we again explore the relationship between the shadow value of external finance and the two variables found to be significant in the financial constraint,

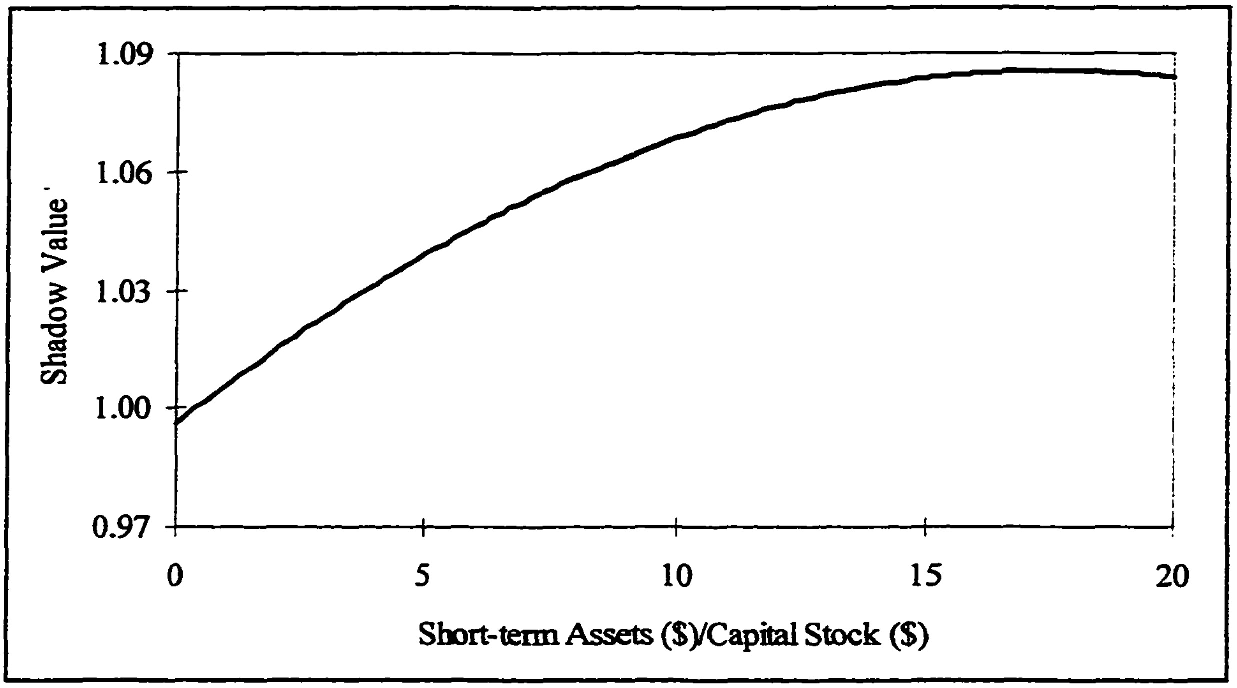


Figure 7.3. Change in the shadow value of external finance due to short-term asset value

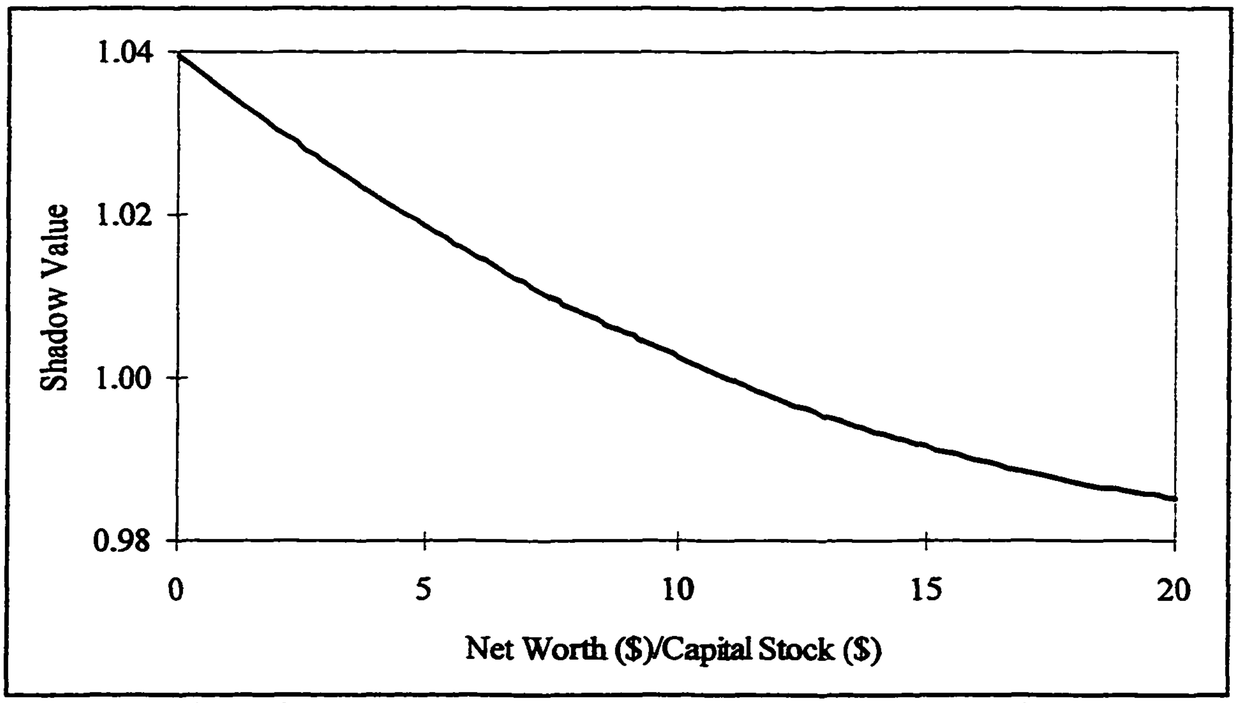


Figure 7.4. Change in the shadow value of external finance due to net worth

net worth and the value of short-term assets. As with the previous graphs, they are based on the parameter estimates from Model 3 with all other variables set at their mean values. Unlike before, the quadratic terms have a visible impact in these relationships, but the basic relationships remain the same. As the value of short-term assets increases, the shadow value of finance increases at a decreasing rate. As the ratio of net worth to the capital stock increases, the shadow value of finance decreases at a decreasing rate.

7.3 Examining Reduced Models

Based on Model 3 from the gross investment analysis, we now explore reducing the variables contained in the financial constraint. One reason to explore this is the possibility of strong multicollinearity among the financial variables. When regressors are highly correlated, we often see three problems that occur during the estimation procedure. These are:

1) parameter estimates may have large standard errors and low significance levels when they are jointly significant and the fit of the model is quite good;
2) small changes in the data produce large changes in parameter estimates; and
3) parameter estimates may have the wrong signs or implausible magnitudes

(Greene, 1990).

When we compare Models 3 and 4 from the gross investment results (Table 7.2), three of the terms in the financial constraint are statistically significant in Model 3, while none are in Model 4. But both models are not rejected by the J-test. These results suggest that multicollinearity could be a major problem in this analysis. One way to alleviate the multicollinearity problem is to reduce the number of variables in the model; this is the approach we employ. Other approaches include the use of additional information in the model and the use of other estimators, such as ridge regressions and principal component estimation for regression analysis.

We construct the 30 various submodels contained within Model 3 by omitting one or more of the variables from the financial constraint. When we omit a variable, both the linear and quadratic terms are removed. Table 7.3 lists the 30 submodels and shows the J- test results of each submodel versus Model 3 from Table 7.2. A p-value above 0.05 indicates evidence for accepting the reduced model over Model 3. The parameter estimates for all of these submodels are given in Appendix 3.

Table 7.3. Testing the submodels

Variables	χ^2	d. f.	p-value	Variables	χ^2	d. f.	p-value
V, TL, NW, CL	1.84	2	0.3991	NW, CL	14.31	6	0.0264
Age, TL, NW, CL	5.20	2	0.0742	TL, CL	10.77	6	0.0958
Age, V, NW, CL	3.19	2	0.2029	TL, NW	12.94	6	0.0439
Age, V, TL, CL	6.15	2	0.0461	V, CL	DNC	6	
Age, V, TL, NW	4.20	2	0.1226	V, NW	11.44	6	0.0756
TL, NW, CL	10.76	4	0.0294	V, TL	DNC	6	
V, NW, CL	3.80	4	0.4332	Age, CL	DNC	6	
V, TL, CL	10.77	4	0.0293	Age, NW	13.44	6	0.0365
V, TL, NW	11.27	4	0.0237	Age, TL	10.00	6	0.1248
Age, NW, CL	8.56	4	0.0732	Age, V	DNC	6	
Age, TL, CL	6.88	4	0.1424	CL	19.47	8	0.0125
Age, TL, NW	6.12	4	0.1903	NW	14.14	8	0.0781
Age, V, CL	DNC*	4		TL	14.08	8	0.0797
Age, V, NW	6.77	4	0.1486	V	DNC	8	
Age, V, TL	3.00	4	0.5582	Age	17.94	8	0.0217

*DNC stands for did not converge.

Fifteen of the submodels are preferred to the full model, Model 3. To quickly summarize the patterns of the submodel estimates, we concentrate on three of the submodels. The parameter estimates from these submodels are given in Table 7.4. Many of the other submodels have results which are very similar to these three submodels. The first submodel has statistically significant parameter estimates for three of the financial variables and a small adjustment cost parameter estimate. Four of the 15 submodels show such a pattern with at least one of the financial variables having a significant parameter estimate.

Table 7.4. Selected submodel results

Parameter	Value	S. E.	Value	S. E.	Value	S. E.
Adj. costs						
θ_0	0.015	0.052	0.127^*	0.051	-1.89^*	0.74
Fin. Constraint						
ρ_0	0.035	0.025	0.034	0.021	-13	31
ρ_1 (AGE_{t-1})			-0.0042	0.0025	-0.03	0.15
ρ_2 (V_t)	0.0046^*	0.0023			5	10
ρ_3 (TL_{t-1})			0.0026	0.0017	0.14	0.28
ρ_4 (NW_{t-1})	-0.0019	0.0015	-0.0006	0.0013		
ρ_5 (CL_{t-1})	0.0107	0.0059				
ρ_6 $(AGE_{t-1})^2$			$-2.5\text{x}10^{-4}$	$1.3\text{x}10^{-4}$	-0.010	0.016
ρ_7 $(V_t)^2$	$-1.98\text{x}10^{-4*}$	$7.9\text{x}10^{-5}$			-0.37	0.82
ρ_8 $(TL_{t-1})^2$			$-4.8\text{x}10^{-5}$	$3.9\text{x}10^{-5}$	-0.007	0.016
ρ_9 $(NW_{t-1})^2$	$8.0\text{x}10^{-5*}$	$3.3\text{x}10^{-5}$	$7.1\text{x}10^{-6}$	$9.6\text{x}10^{-6}$		
ρ_{10} $(CL_{t-1})^2$	$-1.80\text{x}10^{-3*}$	$7.1\text{x}10^{-4}$				
J-stat. tests[a]		p-value		p-value		p-value
χ^2 (O. R.)	8.53	0.5774	10.85	0.3696	7.72	0.6559
d. f.	10		10		10	
χ^2 (vs. full)	3.80	0.4332	6.12	0.1903	3.00	0.5582
d. f.	4		4		4	

[a]O. R. stands for overidentifying restrictions, full stands for Model 3 in Table 7.2. The parameter estimates for the time effects are not presented.

*Significantly different from zero at the 5% level based on the two-tailed *t*-statistic.

Ten of the submodels have parameter estimates similar to the second submodel in Table 7.4. The adjustment cost parameter estimates is twice as large as the same estimate for the full model. If anything is statistically significant from the financial constraint, it is the intercept. The third submodel is unique in that it is the only submodel preferred over the full model to have a negative adjustment cost parameter estimate and the estimate is statistically significant. None of the parameter estimates from the financial constraint are significant, but there is an abrupt change in the parameter estimates. Whereas all of the other submodels and the full models have mean shadow values of external finance around 100 percent, this submodel has a mean shadow value of zero percent.

7.4 Possible Reasons for Results

The estimations from the Euler equation specification contain a mixture of positive and negative results, mostly negative. Some of the estimations indicate that a financial constraint is relevant and net worth, the value of short-term assets, and current liabilities are significant. However, the results are not robust. Estimates for the shadow value of external finance often exceed 100 percent. Adjustment costs are estimated either to be negative or extremely small. Several problems seem to plague the analysis.

One possible problem we examine is the inclusion of 1993 in our short panel data set. Agriculture in Iowa suffered a great deal in that year due to extreme wet conditions in the spring and the floods of that summer. These weather events put a strong financial burden on many farms and may have moved farmers to make unusual decisions. To explore whether

1993 had adversely impacted the results, we estimate the traditional Euler equation with symmetric adjustment costs and no financial constraint on a year-by-year basis. The estimates are very similar for each of the years, indicating 1993 did not adversely impact the study.

The adjustment cost functions we employed in the model may not be rich enough to capture investment adjustment costs. Several studies have explored adjustment cost issues. Goolsbee and Gross (1997) examined adjustment costs for airlines and find non-convexities in adjustment costs at the plant level. Firms have a large area of investment inactivity where desired and actual output may differ between 10 and 40 percent before investing. Hamermesh and Pfann (1996) reviewed various adjustment cost specifications and point out that although some firms may face symmetric adjustment costs, often this specification is dominated by some other specification. Most microeconomic data do not support the symmetric convex adjustment cost assumption. Our results are indicative of this last statement.

In another attempt, we modify the adjustment cost function to another asymmetric cost form given by

$$A(I_t, K_{t-1}) = \left[0.5\theta_0\left(\frac{I_t}{K_{t-1}} - \upsilon\right)^2 - \theta_1\frac{I}{K_{t-1}} + \exp\left(\theta_1\frac{I}{K_{t-1}}\right) - 1\right]K_{t-1}.$$

If θ_1 is zero, adjustment costs are symmetric. If θ_1 is positive, marginal adjustment costs are higher for positive investment than for disinvestment. Marginal adjustment costs for disinvestment are higher than for positive investment if θ_1 is negative. Parameter estimation with this form of adjustment costs ended with either non-convergence of the estimation or implausible parameter estimates.

Within the last few years, some studies have questioned both the Euler equation approach and the GMM estimation technique. Carroll (1997) investigated consumption Euler equation estimation. His analysis suggested that the application of GMM to the full nonlinear Euler equation can suffer greatly if there is measurement error in the data. In a series of papers, Oliner, Rudebusch, and Sichel (1995, 1996) attacked the GMM-Euler equation combination. They found that more traditional models of investment outperform Euler equation models in forecasting investment and that Euler equation models have much larger squared forecast errors. They also found the parameter estimates from Euler equations models to display instability.

Examining inventory from manufacturing data, Fuhrer, Moore, and Schuh (1995) compared GMM and maximum likelihood (ML) parameter estimates. Their findings indicated the GMM estimates are "often biased (apparently due to poor instruments), statistically insignificant, economically implausible, and dynamically unstable." [1] Meanwhile, the ML estimates are generally the opposite. In Monte Carlo simulations, the authors showed GMM to suffer from small sample parameter bias and they related this to the quality of the instrumental variables. Nelson and Startz (1990) cautioned that instrument variables approaches suffer when the instruments are weakly correlated with the explanatory variables.

[1] From the abstract of their paper.

CHAPTER 8. CONCLUSIONS

Using two different approaches, we have examined the relationship between a firm's investment and its financial variables. Under perfect capital markets, there would be no relationship between the two. However, imperfections in the market such as asymmetric information have led researchers to explore these potential relationships. Our study continues in that vein. The 5 Cs of lending (character, capacity, collateral, credit rating, and capital) summarize the attributes lenders desire in borrowers. In both approaches, we incorporated proxies for these characteristics into the farmer's investment decision and explored the impacts of these variables on farm machinery investment.

The first approach consisted of a composite regression model constructed from various elements of traditional investment models and variables representing the 5 Cs of lending. This approach expanded the literature in three ways. First, the inclusion of several financial variables allowed for multiple linkages between investment and financial variables instead of narrowing the focus to one variable. Second, the parameters in the model were estimated using a Bayesian approach which, to our knowledge, has not been employed before in this area. Third, the model we fitted to the data was an extension of the usual mixed linear model, where the distribution of the residuals was taken to be a mixture of normal distributions with unknown mixing proportions and unknown variance components. We used a stochastic variable selection approach based on Bayes factors to select the fixed regressors in the model.

The second approach derived an investment equation from the firm's optimization problem, an Euler equation approach. The 5 Cs of lending were incorporated into the problem through a multiplier associated with a borrowing constraint. This approach also extends the literature through the possibility of multiple links between investment and financial variables. Also, our study is one of a very limited number of agricultural investment studies (to our knowledge, there is only one other study) to use farm level data with the Euler equation approach.

The data set is composed of 590 Iowa farms that are members of the Iowa Farm Business Association and have reported farm level financial and production data from 1991 to 1995. The use of farm level data has several advantages in this type of study. The theory underlying investment models is based on firm level decisions. Most investment studies, though, have estimated investment models on aggregate industry data. The credit constraint issue is also an area where the theory behind the models originates from firm level decisions. Our study employs data that are at the level at which the theory is developed.

For both approaches we incorporate the 5 Cs of lending through the use of proxy variables. The five variables chosen to represent the 5 Cs are operator age, the value of short-term assets, total farm liabilities, net worth, and current farm liabilities. For the composite regression analysis, we begin with estimates from the classical mixed model approach. The regression model combines the variables listed above with elements of other traditional investment models, such as the change of output from the accelerator model. Of the nine variables included in the analysis, all but one (total liabilities) have terms with statistically significant parameter estimates. Elasticities indicate that operator age, the cost of capital, and the value of short-term assets have the largest relative impacts on farm machinery investment.

We considered various formulations of the model described earlier when performing the estimation within the Bayesian framework. The first model formulation we examined included the variable selection component in the estimation but not the outlier detection component. This set-up is the most closely related to the classical mixed model in which variables were included in the model without any selection procedure and errors were taken to be normally distributed. Only six of the nine regressors are selected over 80 percent of the time. The three variables that are not selected are total farm liabilities, net worth, and current farm liabilities, three of the proxies for the 5 Cs of lending. However, many of the results from this formulation parallel those from the classical mixed model.

The next model formulation adds the outlier detection component to the model. Change in output, the value of short-term assets, lagged machinery value, and lagged investment are selected at least 80 percent of the time. Thus, only one of the variables representing the 5 Cs of lending is supported under the model. The variable from the accelerator model of investment, change in output, is also supported; but the neoclassical variable, the cost of capital, is not. The outlier detection component indicates that 25 to 30 percent of the observations may be outliers. The addition of the outlier detection component has a strong impact on the error variance estimates as expected. The estimate of the residual variance obtained from a classical viewpoint and from the Bayesian approach without the outlier detection component was equal to about five. With the outlier detection component, the error variance is estimated to be less than one. Elasticity measures based on the parameter estimates from this simulation show lagged machinery value, the value of short-term assets, and operator age to have the largest relative impacts on investment.

Eight other model formulations, obtained by changing the settings of the variable selection and outlier detection components, were also considered, and marginal posterior distributions for all parameters were obtained. The results are consistent with those obtained from the original model formulation. As the prior probability of variable inclusion is changed from 50 to 90 percent, operator age and lagged investment are added to the selected variable list. As the prior probability of variable inclusion is changed from 50 to 10 percent, change in output is removed from the selected variable list. Posterior means, variance estimates, and the outlier detection hyperparameter estimates are all quite consistent across the alternative model formulations.

The results imply strong support for the accelerator model of investment with the inclusion of other relevant variables, lagged machinery value, lagged investment, and the value of short-term assets (one of the proxies for the 5 Cs). Another one of the proxies, operator age, receives less support. The other proxies, net worth and the liability measures, receive little to no support, which is unexpected since these variables are usually among the first financial variables researchers add to investment models.

The Bayesian framework with the variable selection and outlier detection components works very well. This structure could be put in place to examine many issues in agricultural economics and many other fields. Future research efforts include the application and extension of this type of model. For example, estimating production technical efficiency is one area in which this approach may be useful.

The Euler equation approach is more problematic. Under the original specification looking at net investment, all models are rejected and the preferred model is also the most restrictive, with symmetric adjustment costs and no financial constraint. Within the financial constraint, only the value of short-term assets and net worth are ever found to be statistically significant. Estimated adjustment costs are either negative or positive but very small. The shadow value of external finance is estimated to be around 100 percent.

One possible explanation for these results can be found from the examination of net, versus gross, investment with measurement errors in the value of the capital stock. To explore this issue, we also estimate the models for gross investment. Three of the four models are not rejected by Hansen's J-test. The preferred model has symmetric adjustment costs and a financial constraint. The effects for the value of short-term assets and net worth are significantly different from zero. But many of the problems that occurred in the net investment analysis also occur in the gross investment analysis. For both of the Euler equation analyses, higher shadow values of external finance are linked to higher values of short-term assets and lower values of net worth.

Multicollinearity is a strong possibility with the financial variables included in this analysis. We examine the 30 various submodels contained within the financial constraint specification to monitor whether multicollinearity affected the results and to see if any of the reduced forms would be preferred over the original financial constraint. Fifteen of the submodels are preferred over the original, but many of these have no statistically significant parameter values.

Given the mostly negative results from the Euler equation framework, we review the possible reasons for them. The data set has a small time series component and one of the years could be considered extreme. The data do not support the econometric models on adjustment costs, as evidenced by the many negative parameter estimates. Recent studies have also found several weaknesses in the Euler equation - generalized method of moments combination. More traditional models of investment outperform Euler equation models in forecasting. Parameter estimates display instability. GMM has also been found to suffer from small sample bias.

In his study of consumption Euler equations, Carroll (1997) suggested that the Euler equation approach "should be abandoned" for other econometric approaches. Based on the results from this study, we are more inclined to think in that direction also. Further research needs to address the small samples of GMM estimators. The stability of the parameter estimates are of the utmost importance, especially for analyses done at the aggregate level.

APPENDIX 1. TESTING THE BAYESIAN PROGRAM

Before estimating the various models, we tested our Bayesian computer programs on two simulated data sets. In the first data set, we created 225 observations (15 groups with 15 time periods of data) of ten independent variables from several Normal distributions in Microsoft Excel and saved these as constants. Next we created annual random effects (Normal (0, 1)) and residual errors (Normal (0, 16) except for nine outliers, Normal (0, 1600)). We computed the dependent variable as a linear function of the constants, the random effects, and the residuals. We then modified the model outlined in Chapter 4 for this data and estimated the model twenty times to see how the variable selection and outlier detection components would perform under a controlled environment.

Table A1.1 shows the formula for the dependent variable and the summary statistics for the regressors in this example. Two of the ten independent variables are not used in the

Table A1.1. Summary of the 1st test data set

$$Y = 20 + 1*X1 - 2*X2 + 3*X3 - 5*X5 + 6*X6 - 7*X7 - 9*X9 + 10*X10$$

Variable	Minimum	Mean	Maximum
X1	-11.162	3.852	15.274
X2	5.184	10.029	14.594
X3	-22.755	-6.471	11.884
X4	-7.058	-3.874	-0.798
X5	7.571	15.260	23.020
X6	0.606	2.003	3.398
X7	12.553	25.293	44.197
X8	-28.076	-15.426	-3.782
X9	0.301	0.967	1.630
X10	1.272	6.016	10.949

calculation of the dependent variable. Each of the twenty Bayesian simulation runs consists of four chains of four thousand iterations, for a total of 16,000 observations. In each run, the prior probability for variable selection is set at one-half and the prior for the proportion of outliers is set at 0.1. Behavior of the chains is monitored by Gelman and Rubin's R-statistic, $\sqrt{\hat{R}}$. Convergence is assumed when $\sqrt{\hat{R}}$ is below 1.2 for all parameters.

In Table A1.2, we summarize the results of the twenty simulations. The simulations went extremely well. The means for the parameter estimates are nearly identical to the actual values. The variable selection component correctly chose the eight variables included in the equation. The outlier detection component correctly identified the nine outliers at least 85 percent of the time. The other observations were identified as outliers less than eight percent of the time.

Table A1.2. Summary of the results for the 1st test data set

		Posterior quantiles					% of times
Variable	Mean	2.5%	25%	50%	75%	97.5%	chosen
X1 (1)*	0.996	0.839	0.943	0.996	1.050	1.153	99.96
X2 (-2)	-2.014	-2.348	-2.128	-2.013	-1.901	-1.685	99.80
X3 (3)	3.002	2.879	2.960	3.002	3.044	3.126	100.00
X4 (0)	0.000	0.000	0.000	0.000	0.000	0.000	1.45
X5 (-5)	-5.007	-5.228	-5.082	-5.005	-4.930	-4.793	100.00
X6 (6)	5.996	4.730	5.559	5.993	6.435	7.266	99.62
X7 (-7)	-6.977	-7.087	-7.015	-6.977	-6.939	-6.866	100.00
X8 (0)	0.000	0.000	0.000	0.000	0.000	0.000	0.89
X9 (-9)	-8.851	-11.184	-9.644	-8.843	-8.052	-6.549	99.43
X10 (10)	10.015	9.689	9.902	10.015	10.127	10.340	100.00
Parameter							
η (0.96)	0.927	0.875	0.912	0.929	0.943	0.965	
σ_y^2 (1)	0.610	0.209	0.372	0.520	0.745	1.538	
σ_ε^2 (16)	19.705	15.753	18.166	19.567	21.098	24.391	

*Numbers in parenthesis are the actual values.

For the second test, we manufactured an investment data set based on the actual investment data we are studying. Our model contains 54 effects (9 main linear effects, 9 quadratic effects, and 36 cross effects). Ten of these were chosen at random and combined to form a series of hypothetical investment data. Thus the hypothetical data set would have the same structure as the actual data set, but we would precisely know the data generating mechanism for investment. We then estimated the parameters in the model as we would with the actual data set with our full model specification under the variable selection and outlier detection components. The equation for investment is given by:

$$\begin{aligned} I^N_{i,t} = \Delta Q_{i,t} + AGE_{i,t-1} + 0.1*\Delta Q_{i,t}^2 - 0.01*TL_{i,t-1}^2 + 0.1*K_{i,t-1}^2 - V_{i,t}*I^N_{i,t-1} \\ + 0.1*NW_{i,t-1}*CL_{i,t-1} - 0.01*NW_{i,t-1}*K_{i,t-1} + NW_{i,t-1}*I^N_{i,t-1} \\ - K_{i,t-1}*I^N_{i,t-1} + \varepsilon_{i,t} \end{aligned}$$

where $\varepsilon_{i,t}$ is a standard normal random disturbance. We have not built in an intercept, any annual random effects, or any outliers; but our estimation procedure will search for and include these features.

For this test, we simulated four chains with 1,000 iterations each. Convergence is monitored by Gelman and Rubin's R-statistic, $\sqrt{\hat{R}}$. Table A1.3 summarizes the results of the simulations. The simulations were long enough to allow most of the estimates to be considered "converged" with exceptions being the intercept and annual random effects. The variable selection component performed rather well. It selected the correct variables a vast majority of the time. While each of the included variables was chosen at least 98.7 percent of the time, each of the excluded variables was chosen for the model less than 16.2 percent of the time with most of those below five percent. Due to the structure of our model for the

Table A1.3. Summary of the results for the 2nd test data set

Variable	Mean	Posterior quantiles 2.5%	50%	97.5%	$\sqrt{\hat{R}}$	% of times chosen
Intercept	0.22	-0.51	0.01	1.29	3.60	100.00
$\Delta Q_{i,t}$	1.00	1.00	1.00	1.01	1.00	100.00
$V_{i,t}$	0.00	0.00	0.00	0.00	1.19	5.75
$C_{i,t}$	0.00	0.00	0.00	0.00	1.10	5.58
$AGE_{i,t-1}$	1.00	1.00	1.00	1.00	1.00	100.00
$TL_{i,t-1}$	0.00	0.00	0.00	0.00	1.00	11.35
$NW_{i,t-1}$	0.00	0.00	0.00	0.00	1.00	2.53
$CL_{i,t-1}$	0.00	0.00	0.00	0.00	1.01	9.40
$K_{i,t-1}$	0.00	0.00	0.00	0.00	1.02	8.70
$I^N_{i,t-1}$	-0.00	0.00	0.00	0.00	1.04	3.13
$\Delta Q_{i,t}^2$	0.10	0.10	0.10	0.10	1.00	100.00
$V_{i,t}^2$	0.00	0.00	0.00	0.00	1.00	5.68
$C_{i,t}^2$	-0.00	0.00	0.00	0.00	1.07	1.40
$AGE_{i,t-1}^2$	0.00	0.00	0.00	0.00	1.00	0.33
$TL_{i,t-1}^2$	-0.01	-0.01	-0.01	-0.01	1.00	99.48
$NW_{i,t-1}^2$	0.00	0.00	0.00	0.00	1.00	6.60
$CL_{i,t-1}^2$	0.00	0.00	0.00	0.00	1.00	16.15
$K_{i,t-1}^2$	0.10	0.10	0.10	0.10	1.01	99.80
$I^{N\,2}_{i,t-1}$	0.00	0.00	0.00	0.00	1.01	3.40
$\Delta Q_{i,t}*V_{i,t}$	0.00	0.00	0.00	0.00	1.00	1.73
$\Delta Q_{i,t}*C_{i,t}$	0.00	0.00	0.00	0.00	1.00	0.93
$\Delta Q_{i,t}*AGE_{i,t-1}$	0.00	0.00	0.00	0.00	1.00	0.20
$\Delta Q_{i,t}*TL_{i,t-1}$	0.00	0.00	0.00	0.00	1.00	0.75
$\Delta Q_{i,t}*NW_{i,t-1}$	0.00	0.00	0.00	0.00	1.00	0.53
$\Delta Q_{i,t}*CL_{i,t-1}$	0.00	0.00	0.00	0.00	1.00	0.68
$\Delta Q_{i,t}*K_{i,t-1}$	0.00	0.00	0.00	0.00	1.00	0.15
$\Delta Q_{i,t}*I^N_{i,t-1}$	-0.00	0.00	0.00	0.00	1.12	1.53
$V_{i,t}*C_{i,t}$	0.00	0.00	0.00	0.00	1.00	0.50
$V_{i,t}*AGE_{i,t-1}$	0.00	0.00	0.00	0.00	1.00	3.43
$V_{i,t}*TL_{i,t-1}$	0.00	0.00	0.00	0.00	1.00	11.65
$V_{i,t}*NW_{i,t-1}$	0.00	0.00	0.00	0.00	1.00	7.03
$V_{i,t}*CL_{i,t-1}$	0.00	0.00	0.00	0.00	1.00	13.50
$V_{i,t}*K_{i,t-1}$	0.00	0.00	0.00	0.00	1.00	6.05
$V_{i,t}*I^N_{i,t-1}$	-1.00	-1.00	-1.00	-1.00	1.00	99.88
$C_{i,t}*AGE_{i,t-1}$	0.00	0.00	0.00	0.00	1.00	0.33
$C_{i,t}*TL_{i,t-1}$	0.00	0.00	0.00	0.00	1.00	0.33
$C_{i,t}*NW_{i,t-1}$	0.00	0.00	0.00	0.00	1.00	0.30
$C_{i,t}*CL_{i,t-1}$	0.00	0.00	0.00	0.00	1.00	0.30

Table A1.3. (continued)

Variable	Mean	Posterior quantiles 2.5%	50%	97.5%	$\sqrt{\hat{R}}$	% of times chosen
$C_{i,t}*K_{i,t-1}$	0.00	0.00	0.00	0.00	1.33	0.63
$C_{i,t}*I^{N}_{i,t-1}$	-0.00	0.00	0.00	0.00	1.01	1.88
$AGE_{i,t-1}*TL_{i,t-1}$	0.00	0.00	0.00	0.00	1.00	6.63
$AGE_{i,t-1}*NW_{i,t-1}$	0.00	0.00	0.00	0.00	1.00	4.75
$AGE_{i,t-1}*CL_{i,t-1}$	0.00	0.00	0.00	0.00	1.00	9.80
$AGE_{i,t-1}*K_{i,t-1}$	0.00	0.00	0.00	0.00	1.00	4.13
$AGE_{i,t-1}*I^{N}_{i,t-1}$	0.00	0.00	0.00	0.00	1.00	3.03
$TL_{i,t-1}*NW_{i,t-1}$	0.00	0.00	0.00	0.00	1.00	11.45
$TL_{i,t-1}*CL_{i,t-1}$	0.00	0.00	0.00	0.00	1.00	8.25
$TL_{i,t-1}*K_{i,t-1}$	0.00	0.00	0.00	0.00	1.00	7.25
$TL_{i,t-1}*I^{N}_{i,t-1}$	-0.00	0.00	0.00	0.00	1.01	6.48
$NW_{i,t-1}*CL_{i,t-1}$	0.10	0.10	0.10	0.10	1.00	100.00
$NW_{i,t-1}*K_{i,t-1}$	-0.01	-0.01	-0.01	-0.01	1.00	98.70
$NW_{i,t-1}*I^{N}_{i,t-1}$	1.00	1.00	1.00	1.00	1.00	100.00
$CL_{i,t-1}*K_{i,t-1}$	0.00	0.00	0.00	0.00	1.00	9.60
$CL_{i,t-1}*I^{N}_{i,t-1}$	0.00	0.00	0.00	0.00	1.00	6.25
$K_{i,t-1}*I^{N}_{i,t-1}$	-1.00	-1.00	-1.00	-1.00	1.00	99.95
Parameter						
η	1.00	0.99	1.00	1.00	1.01	
σ_{ε}^{2}	0.99	0.92	0.99	1.06	1.00	
σ_{γ}^{2}	0.48	0.26	0.44	0.93	1.13	

actual data set, an intercept and annual random effects are always simulated. Since the fabricated data employed here did not include an intercept or annual random effects, it is not surprising that the model struggled to handle these factors and basically set the intercept to cancel the annual effects.

The mean parameter estimates are nearly identical to the actual values and the spread of the estimates, as shown by the quantiles from the simulation sample, is quite narrow around the means. The outlier detection component also performed quite well. Only 12 of the observations (less than one percent of the observations) were chosen as outliers over ten percent of the time. The error variance mean estimate is also very near the actual value. The table includes the estimates for the annual random effects and their variance. Given the results of this test of our program, we proceeded with the estimation with actual data.

APPENDIX 2. POSTERIOR DISTRIBUTIONS FROM DIFFERENT MODEL FORMULATIONS

Table A2.1. Summary of results for model Var1Out1

	Posterior	Posterior quantiles				% of times
Variable	Mean	2.5%	50%	97.5%	$\sqrt{\hat{R}}$	chosen
Intercept	-0.039	-0.909	-0.019	0.766	1.04	100.00
$\Delta Q_{i,t}$	0.009	0.000	0.000	0.036	1.00	40.45
$V_{i,t}$	0.017	0.009	0.017	0.025	1.00	98.29
$C_{i,t}$	0.005	0.000	0.000	0.093	1.05	6.38
$AGE_{i,t-1}$	-0.001	-0.012	0.000	0.000	1.00	15.34
$TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.04
$NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.03
$CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.02	0.10
$K_{i,t-1}$	-0.061	-0.074	-0.061	-0.046	1.00	99.97
$I^N_{i,t-1}$	-0.008	-0.074	0.000	0.000	1.01	14.19
$\Delta Q_{i,t}^2$	0.000	0.000	0.000	0.000	1.00	0.01
$V_{i,t}^2$	0.000	0.000	0.000	0.000	1.00	0.20
$C_{i,t}^2$	0.000	0.000	0.000	0.000	1.07	0.23
$AGE_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.01
$TL_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.01
$NW_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.01
$CL_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.03
$K_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.02
$I^N{}_{i,t-1}^2$	0.017	0.000	0.019	0.025	1.03	87.33
$\Delta Q_{i,t}*V_{i,t}$	0.000	0.000	0.000	0.000	1.00	0.01
$\Delta Q_{i,t}*C_{i,t}$	0.000	0.000	0.000	0.000	1.00	0.03
$\Delta Q_{i,t}*AGE_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.02
$\Delta Q_{i,t}*TL_{i,t-1}$	0.000	-0.001	0.000	0.000	1.01	6.27
$\Delta Q_{i,t}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$\Delta Q_{i,t}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.01	2.27
$\Delta Q_{i,t}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.21
$\Delta Q_{i,t}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.05
$V_{i,t}*C_{i,t}$	0.000	0.000	0.000	0.000	1.00	0.03
$V_{i,t}*AGE_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$V_{i,t}*TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.02
$V_{i,t}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.40
$V_{i,t}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.08
$V_{i,t}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.04
$V_{i,t}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.41	0.08
$C_{i,t}*AGE_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.03
$C_{i,t}*TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.02	0.54
$C_{i,t}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.02

Table A2.1. (continued)

	Posterior	Posterior quantiles				% of times
Variable	Mean	2.5%	50%	97.5%	$\sqrt{\hat{R}}$	chosen
$C_{i,t}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.09	0.23
$C_{i,t}*K_{i,t-1}$	-0.001	-0.010	0.000	0.000	1.00	6.88
$C_{i,t}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.01	0.63
$AGE_{i,t-1}*TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$AGE_{i,t-1}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$AGE_{i,t-1}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$AGE_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.02
$AGE_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.09	0.11
$TL_{i,t-1}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.04
$TL_{i,t-1}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$TL_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$TL_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.02
$NW_{i,t-1}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.05
$NW_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$NW_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.03
$CL_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$CL_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.03	0.33
$K_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.05	1.37
Parameter						
η	0.741	0.703	0.741	0.778	1.00	
σ_ε^2	0.939	0.832	0.937	1.061	1.00	
σ_v^2	0.458	0.248	0.429	0.834	1.00	

Table A2.2. Summary of results for model Var1Out5

	Posterior	Posterior quantiles				% of times
Variable	Mean	2.5%	50%	97.5%	$\sqrt{\hat{R}}$	chosen
Intercept	-0.182	-1.110	-0.201	0.871	1.04	100.00
$\Delta Q_{i,t}$	0.010	0.000	0.000	0.039	1.01	43.72
$V_{i,t}$	0.017	0.007	0.017	0.025	1.01	97.98
$C_{i,t}$	0.007	0.000	0.000	0.111	1.02	6.69
$AGE_{i,t-1}$	-0.002	-0.012	0.000	0.000	1.00	15.80
$TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.05
$NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.02
$CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.09	0.07
$K_{i,t-1}$	-0.060	-0.074	-0.060	-0.044	1.01	99.95
$I^N_{i,t-1}$	-0.008	-0.073	0.000	0.000	1.00	13.31
$\Delta Q_{i,t}^2$	0.000	0.000	0.000	0.000	1.00	0.02
$V_{i,t}^2$	0.000	0.000	0.000	0.000	1.00	0.26
$C_{i,t}^2$	0.000	0.000	0.000	0.000	1.03	0.20
$AGE_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.01
$TL_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.01
$NW_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.01
$CL_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.02
$K_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.02
$I^{N\,2}_{i,t-1}$	0.016	0.000	0.018	0.025	1.03	83.68
$\Delta Q_{i,t}*V_{i,t}$	0.000	0.000	0.000	0.000	1.00	0.02
$\Delta Q_{i,t}*C_{i,t}$	0.000	0.000	0.000	0.000	1.00	0.03
$\Delta Q_{i,t}*AGE_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$\Delta Q_{i,t}*TL_{i,t-1}$	0.000	-0.001	0.000	0.000	1.06	6.02
$\Delta Q_{i,t}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$\Delta Q_{i,t}*CL_{i,t-1}$	0.000	-0.001	0.000	0.000	1.05	3.04
$\Delta Q_{i,t}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.01	0.24
$\Delta Q_{i,t}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.04
$V_{i,t}*C_{i,t}$	0.000	0.000	0.000	0.000	1.00	0.04
$V_{i,t}*AGE_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.02
$V_{i,t}*TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.03
$V_{i,t}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.26
$V_{i,t}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.07
$V_{i,t}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.03
$V_{i,t}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.24	0.13
$C_{i,t}*AGE_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.03
$C_{i,t}*TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.03	0.44
$C_{i,t}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$C_{i,t}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.11	0.27

Table A2.2. (continued)

	Posterior	Posterior quantiles				% of times
Variable	Mean	2.5%	50%	97.5%	$\sqrt{\hat{R}}$	chosen
$C_{i,t}*K_{i,t-1}$	-0.001	-0.010	0.000	0.000	1.00	6.19
$C_{i,t}*I^{N}_{i,t-1}$	0.000	0.000	0.000	0.000	1.01	0.54
$AGE_{i,t-1}*TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$AGE_{i,t-1}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$AGE_{i,t-1}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$AGE_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.02
$AGE_{i,t-1}*I^{N}_{i,t-1}$	0.000	0.000	0.000	0.000	1.03	0.10
$TL_{i,t-1}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.02
$TL_{i,t-1}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$TL_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$TL_{i,t-1}*I^{N}_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.03
$NW_{i,t-1}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.04
$NW_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$NW_{i,t-1}*I^{N}_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$CL_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.02
$CL_{i,t-1}*I^{N}_{i,t-1}$	0.000	0.000	0.000	0.000	1.19	0.47
$K_{i,t-1}*I^{N}_{i,t-1}$	0.000	0.000	0.000	0.003	1.04	2.63
Parameter						
η	0.732	0.693	0.732	0.771	1.00	
σ_{ε}^{2}	0.921	0.813	0.919	1.040	1.00	
σ_{y}^{2}	0.469	0.251	0.439	0.864	1.01	

Table A2.3. Summary of results for model Var1Out9

Variable	Posterior Mean	Posterior quantiles 2.5%	50%	97.5%	$\sqrt{\hat{R}}$	% of times chosen
Intercept	-0.104	-0.946	-0.084	0.612	1.12	100.00
$\Delta Q_{i,t}$	0.012	0.000	0.000	0.039	1.01	46.99
$V_{i,t}$	0.018	0.010	0.018	0.026	1.00	99.19
$C_{i,t}$	0.005	0.000	0.000	0.088	1.02	4.58
$AGE_{i,t-1}$	-0.001	-0.012	0.000	0.000	1.00	15.36
$TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.04
$NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.03
$CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.21	0.08
$K_{i,t-1}$	-0.061	-0.075	-0.061	-0.045	1.00	99.96
$I^N_{i,t-1}$	-0.008	-0.074	0.000	0.000	1.00	13.01
$\Delta Q_{i,t}^2$	0.000	0.000	0.000	0.000	1.00	0.01
$V_{i,t}^2$	0.000	0.000	0.000	0.000	1.00	0.13
$C_{i,t}^2$	0.000	0.000	0.000	0.000	1.12	0.17
$AGE_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.01
$TL_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.01
$NW_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.01
$CL_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.04
$K_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.01
${I^N_{i,t-1}}^2$	0.015	0.000	0.018	0.025	1.01	83.58
$\Delta Q_{i,t}*V_{i,t}$	0.000	0.000	0.000	0.000	1.00	0.03
$\Delta Q_{i,t}*C_{i,t}$	0.000	0.000	0.000	0.000	1.00	0.05
$\Delta Q_{i,t}*AGE_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$\Delta Q_{i,t}*TL_{i,t-1}$	0.000	-0.001	0.000	0.000	1.03	6.57
$\Delta Q_{i,t}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$\Delta Q_{i,t}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.03	2.11
$\Delta Q_{i,t}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.15
$\Delta Q_{i,t}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.06
$V_{i,t}*C_{i,t}$	0.000	0.000	0.000	0.000	1.00	0.05
$V_{i,t}*AGE_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$V_{i,t}*TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.02
$V_{i,t}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.20
$V_{i,t}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.06
$V_{i,t}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.02
$V_{i,t}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.05	0.11
$C_{i,t}*AGE_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.02
$C_{i,t}*TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.01	0.52
$C_{i,t}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$C_{i,t}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.06	0.25

Table A2.3. (continued)

Variable	Posterior Mean	Posterior quantiles 2.5%	50%	97.5%	$\sqrt{\hat{R}}$	% of times chosen
$C_{i,t}*K_{i,t-1}$	0.000	-0.009	0.000	0.000	1.00	5.33
$C_{i,t}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.02	0.48
$AGE_{i,t-1}*TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$AGE_{i,t-1}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$AGE_{i,t-1}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$AGE_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$AGE_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.04	0.14
$TL_{i,t-1}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$TL_{i,t-1}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$TL_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.02
$TL_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.03
$NW_{i,t-1}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.04
$NW_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$NW_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.02
$CL_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.02
$CL_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.02	0.60
$K_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.01	2.22
Parameter						
η	0.709	0.669	0.709	0.747	1.00	
σ_ε^2	0.879	0.780	0.877	0.991	1.00	
σ_γ^2	0.455	0.249	0.426	0.818	1.00	

Table A2.4. Summary of results for model Var5Out5

Variable	Posterior Mean	Posterior quantiles 2.5%	50%	97.5%	$\sqrt{\hat{R}}$	% of times chosen
Intercept	-0.097	-1.224	-0.058	0.649	1.10	100.00
$\Delta Q_{i,t}$	0.025	0.000	0.026	0.044	1.00	90.21
$V_{i,t}$	0.017	0.010	0.017	0.025	1.00	99.79
$C_{i,t}$	0.008	0.000	0.000	0.092	1.00	14.17
$AGE_{i,t-1}$	-0.004	-0.013	0.000	0.000	1.00	46.85
$TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.26
$NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.19
$CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.03	0.51
$K_{i,t-1}$	-0.060	-0.073	-0.060	-0.046	1.00	100.00
$I^N_{i,t-1}$	-0.040	-0.090	-0.048	0.000	1.00	69.06
$\Delta Q_{i,t}^2$	0.000	0.000	0.000	0.000	1.00	0.06
$V_{i,t}^2$	0.000	0.000	0.000	0.000	1.00	0.54
$C_{i,t}^2$	0.000	0.000	0.000	0.000	1.01	2.75
$AGE_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.05
$TL_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.02
$NW_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.01
$CL_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.11
$K_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.06
$I^N{}_{i,t-1}^2$	0.019	0.000	0.020	0.026	1.02	97.01
$\Delta Q_{i,t}*V_{i,t}$	0.000	0.000	0.000	0.000	1.00	0.16
$\Delta Q_{i,t}*C_{i,t}$	0.000	0.000	0.000	0.000	1.02	0.26
$\Delta Q_{i,t}*AGE_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.06
$\Delta Q_{i,t}*TL_{i,t-1}$	0.000	-0.001	0.000	0.000	1.02	29.66
$\Delta Q_{i,t}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.02
$\Delta Q_{i,t}*CL_{i,t-1}$	0.000	-0.002	0.000	0.000	1.01	15.84
$\Delta Q_{i,t}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.06	1.26
$\Delta Q_{i,t}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.03	0.33
$V_{i,t}*C_{i,t}$	0.000	0.000	0.000	0.000	1.00	0.23
$V_{i,t}*AGE_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.05
$V_{i,t}*TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.06
$V_{i,t}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.87
$V_{i,t}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.21
$V_{i,t}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.10
$V_{i,t}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.24
$C_{i,t}*AGE_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.19
$C_{i,t}*TL_{i,t-1}$	0.000	-0.002	0.000	0.000	1.00	3.45
$C_{i,t}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.06

Table A2.4. (continued)

	Posterior	Posterior quantiles				% of times
Variable	Mean	2.5%	50%	97.5%	$\sqrt{\hat{R}}$	chosen
$C_{i,t}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.01	2.06
$C_{i,t}*K_{i,t-1}$	-0.001	-0.011	0.000	0.000	1.00	18.66
$C_{i,t}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	2.40
$AGE_{i,t-1}*TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.03
$AGE_{i,t-1}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.03
$AGE_{i,t-1}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.06
$AGE_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.06
$AGE_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.81
$TL_{i,t-1}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.07
$TL_{i,t-1}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.03
$TL_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.03
$TL_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.12
$NW_{i,t-1}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.13
$NW_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.03
$NW_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.07
$CL_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.09
$CL_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.01	1.60
$K_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.09	1.84
Parameter						
η	0.732	0.692	0.732	0.769	1.00	
σ_ε^2	0.906	0.804	0.904	1.019	1.00	
σ_y^2	0.458	0.248	0.428	0.844	1.01	

Table A2.5. Summary of results for model Var5Out9

	Posterior	Posterior quantiles				% of times
Variable	Mean	2.5%	50%	97.5%	$\sqrt{\hat{R}}$	chosen
Intercept	-0.149	-1.251	-0.103	0.700	1.06	100.00
$\Delta Q_{i,t}$	0.025	0.000	0.026	0.043	1.00	91.49
$V_{i,t}$	0.018	0.010	0.018	0.026	1.00	99.85
$C_{i,t}$	0.008	0.000	0.000	0.091	1.01	11.93
$AGE_{i,t-1}$	-0.004	-0.013	0.000	0.000	1.00	46.26
$TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.23
$NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.18
$CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.03	0.52
$K_{i,t-1}$	-0.060	-0.074	-0.060	-0.047	1.00	99.99
$I^N_{i,t-1}$	-0.038	-0.091	-0.045	0.000	1.00	67.21
$\Delta Q_{i,t}^2$	0.000	0.000	0.000	0.000	1.00	0.06
$V_{i,t}^2$	0.000	0.000	0.000	0.000	1.00	0.34
$C_{i,t}^2$	0.000	0.000	0.000	0.000	1.02	2.34
$AGE_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.03
$TL_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.01
$NW_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.01
$CL_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.10
$K_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.06
$I^{N\,2}_{i,t-1}$	0.019	0.000	0.020	0.026	1.01	97.90
$\Delta Q_{i,t}*V_{i,t}$	0.000	0.000	0.000	0.000	1.00	0.15
$\Delta Q_{i,t}*C_{i,t}$	0.000	0.000	0.000	0.000	1.04	0.29
$\Delta Q_{i,t}*AGE_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.08
$\Delta Q_{i,t}*TL_{i,t-1}$	0.000	-0.001	0.000	0.000	1.00	28.56
$\Delta Q_{i,t}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.03
$\Delta Q_{i,t}*CL_{i,t-1}$	0.000	-0.002	0.000	0.000	1.03	15.00
$\Delta Q_{i,t}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.01	1.08
$\Delta Q_{i,t}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.05	0.33
$V_{i,t}*C_{i,t}$	0.000	0.000	0.000	0.000	1.00	0.24
$V_{i,t}*AGE_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.06
$V_{i,t}*TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.06
$V_{i,t}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.53
$V_{i,t}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.30
$V_{i,t}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.07
$V_{i,t}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.12	0.24
$C_{i,t}*AGE_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.18
$C_{i,t}*TL_{i,t-1}$	0.000	-0.002	0.000	0.000	1.01	3.86
$C_{i,t}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.08
$C_{i,t}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.01	2.07

Table A2.5. (continued)

	Posterior	Posterior quantiles				% of times
Variable	Mean	2.5%	50%	97.5%	$\sqrt{\hat{R}}$	chosen
$C_{i,t}*K_{i,t-1}$	-0.001	-0.011	0.000	0.000	1.00	15.38
$C_{i,t}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	2.18
$AGE_{i,t-1}*TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.05
$AGE_{i,t-1}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.02
$AGE_{i,t-1}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.06
$AGE_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.07
$AGE_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.01	0.78
$TL_{i,t-1}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.06
$TL_{i,t-1}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.02
$TL_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.04
$TL_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.18
$NW_{i,t-1}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.17
$NW_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.04
$NW_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.08
$CL_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.06
$CL_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.001	1.03	2.91
$K_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.10	1.40
Parameter						
η	0.708	0.668	0.708	0.747	1.00	
σ_ε^2	0.866	0.769	0.864	0.977	1.00	
σ_γ^2	0.465	0.251	0.434	0.860	1.01	

Table A2.6. Summary of results for model Var9Out1

	Posterior	Posterior quantiles				% of times
Variable	Mean	2.5%	50%	97.5%	$\sqrt{\hat{R}}$	chosen
Intercept	-0.046	-0.917	-0.061	0.761	1.08	100.00
$\Delta Q_{i,t}$	0.028	0.011	0.029	0.043	1.00	99.12
$V_{i,t}$	0.017	0.009	0.017	0.026	1.00	99.93
$C_{i,t}$	0.023	-0.007	0.000	0.109	1.01	45.98
$AGE_{i,t-1}$	-0.007	-0.014	-0.008	0.000	1.00	83.19
$TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.01	1.94
$NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	1.36
$CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	4.34
$K_{i,t-1}$	-0.058	-0.071	-0.058	-0.046	1.00	100.00
$I^N_{i,t-1}$	-0.054	-0.094	-0.056	0.000	1.00	94.65
$\Delta Q_{i,t}^2$	0.000	0.000	0.000	0.000	1.00	0.62
$V_{i,t}^2$	0.000	0.000	0.000	0.000	1.00	2.53
$C_{i,t}^2$	-0.004	-0.033	0.000	0.000	1.00	20.18
$AGE_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.27
$TL_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.04
$NW_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.03
$CL_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.73
$K_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.53
$I^{N\,2}_{i,t-1}$	0.020	0.013	0.021	0.027	1.09	99.32
$\Delta Q_{i,t}*V_{i,t}$	0.000	0.000	0.000	0.000	1.00	0.57
$\Delta Q_{i,t}*C_{i,t}$	0.000	0.000	0.000	0.000	1.00	2.54
$\Delta Q_{i,t}*AGE_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.50
$\Delta Q_{i,t}*TL_{i,t-1}$	0.000	-0.001	0.000	0.000	1.01	48.87
$\Delta Q_{i,t}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.28
$\Delta Q_{i,t}*CL_{i,t-1}$	0.000	-0.002	0.000	0.000	1.01	25.23
$\Delta Q_{i,t}*K_{i,t-1}$	0.000	-0.001	0.000	0.000	1.02	3.58
$\Delta Q_{i,t}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.01	2.07
$V_{i,t}*C_{i,t}$	0.000	0.000	0.000	0.000	1.01	2.27
$V_{i,t}*AGE_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.28
$V_{i,t}*TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.26
$V_{i,t}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	3.51
$V_{i,t}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	1.98
$V_{i,t}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.62
$V_{i,t}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	1.65
$C_{i,t}*AGE_{i,t-1}$	0.000	0.000	0.000	0.000	1.01	1.32
$C_{i,t}*TL_{i,t-1}$	0.000	-0.003	0.000	0.000	1.00	9.52
$C_{i,t}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.68
$C_{i,t}*CL_{i,t-1}$	0.000	-0.004	0.000	0.000	1.00	6.21

Table A2.6. (continued)

Variable	Posterior Mean	Posterior quantiles 2.5%	50%	97.5%	$\sqrt{\hat{R}}$	% of times chosen
$C_{i,t}*K_{i,t-1}$	-0.003	-0.012	0.000	0.000	1.00	42.18
$C_{i,t}*I^N_{i,t-1}$	-0.001	-0.017	0.000	0.000	1.00	12.37
$AGE_{i,t-1}*TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.29
$AGE_{i,t-1}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.10
$AGE_{i,t-1}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.62
$AGE_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.94
$AGE_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.002	1.00	4.50
$TL_{i,t-1}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.36
$TL_{i,t-1}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.19
$TL_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.39
$TL_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.05	1.36
$NW_{i,t-1}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.77
$NW_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.11
$NW_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.50
$CL_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.60
$CL_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.006	1.01	8.63
$K_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.001	1.82	2.79
Parameter						
η	0.742	0.703	0.742	0.778	1.00	
σ_ε^2	0.919	0.813	0.917	1.035	1.00	
σ_v^2	0.461	0.251	0.431	0.848	1.00	

Table A2.7. Summary of results for model Var9Out5

	Posterior	Posterior quantiles				% of times
Variable	Mean	2.5%	50%	97.5%	$\sqrt{\hat{R}}$	chosen
Intercept	-0.125	-1.113	-0.145	1.063	1.06	100.00
$\Delta Q_{i,t}$	0.028	0.011	0.029	0.043	1.00	99.00
$V_{i,t}$	0.017	0.009	0.017	0.025	1.01	99.80
$C_{i,t}$	0.023	-0.007	0.000	0.109	1.01	46.13
$AGE_{i,t-1}$	-0.007	-0.014	-0.008	0.000	1.00	83.38
$TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.01	2.11
$NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.03	1.31
$CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.01	4.50
$K_{i,t-1}$	-0.059	-0.072	-0.059	-0.045	1.00	99.99
$I^N_{i,t-1}$	-0.054	-0.093	-0.056	0.000	1.00	94.39
$\Delta Q_{i,t}^2$	0.000	0.000	0.000	0.000	1.00	0.62
$V_{i,t}^2$	0.000	0.000	0.000	0.000	1.00	2.11
$C_{i,t}^2$	-0.003	-0.029	0.000	0.000	1.01	18.99
$AGE_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.20
$TL_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.05
$NW_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.02
$CL_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.86
$K_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.58
$I^{N\,2}_{i,t-1}$	0.020	0.013	0.021	0.026	1.07	99.41
$\Delta Q_{i,t}$*$V_{i,t}$	0.000	0.000	0.000	0.000	1.00	0.63
$\Delta Q_{i,t}$*$C_{i,t}$	0.000	0.000	0.000	0.000	1.00	2.50
$\Delta Q_{i,t}$*$AGE_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.57
$\Delta Q_{i,t}$*$TL_{i,t-1}$	0.000	-0.001	0.000	0.000	1.00	48.04
$\Delta Q_{i,t}$*$NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.17
$\Delta Q_{i,t}$*$CL_{i,t-1}$	0.000	-0.002	0.000	0.000	1.00	24.88
$\Delta Q_{i,t}$*$K_{i,t-1}$	0.000	-0.002	0.000	0.000	1.01	4.01
$\Delta Q_{i,t}$*$I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	2.23
$V_{i,t}$*$C_{i,t}$	0.000	0.000	0.000	0.000	1.02	2.41
$V_{i,t}$*$AGE_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.29
$V_{i,t}$*$TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.27
$V_{i,t}$*$NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	3.31
$V_{i,t}$*$CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	2.28
$V_{i,t}$*$K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.54
$V_{i,t}$*$I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.01	1.61
$C_{i,t}$*$AGE_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	1.42
$C_{i,t}$*$TL_{i,t-1}$	0.000	-0.003	0.000	0.000	1.00	10.16
$C_{i,t}$*$NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.65
$C_{i,t}$*$CL_{i,t-1}$	0.000	-0.005	0.000	0.000	1.01	6.53

Table A2.7. (continued)

Variable	Posterior Mean	Posterior quantiles 2.5%	50%	97.5%	$\sqrt{\hat{R}}$	% of times chosen
$C_{i,t}*K_{i,t-1}$	-0.003	-0.012	0.000	0.000	1.00	41.24
$C_{i,t}*I^N_{i,t-1}$	-0.001	-0.017	0.000	0.000	1.00	12.16
$AGE_{i,t-1}*TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.28
$AGE_{i,t-1}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.10
$AGE_{i,t-1}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.54
$AGE_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.93
$AGE_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.002	1.00	4.49
$TL_{i,t-1}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.36
$TL_{i,t-1}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.16
$TL_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.31
$TL_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.05	1.41
$NW_{i,t-1}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.96
$NW_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.13
$NW_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.54
$CL_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.55
$CL_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.006	1.02	9.55
$K_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.71	2.83
Parameter						
η	0.734	0.695	0.734	0.772	1.00	
σ_ε^2	0.905	0.801	0.903	1.020	1.00	
σ_v^2	0.478	0.254	0.446	0.892	1.01	

Table A2.8. Summary of results for model Var9Out9

	Posterior	Posterior quantiles				% of times
Variable	Mean	2.5%	50%	97.5%	$\sqrt{\hat{R}}$	chosen
Intercept	-0.249	-1.340	-0.253	0.706	1.06	100.00
$\Delta Q_{i,t}$	0.029	0.012	0.030	0.044	1.00	99.17
$V_{i,t}$	0.017	0.009	0.017	0.026	1.00	99.89
$C_{i,t}$	0.022	-0.008	0.000	0.104	1.00	46.18
$AGE_{i,t-1}$	-0.007	-0.014	-0.008	0.000	1.00	83.06
$TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	1.98
$NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.04	1.33
$CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.01	4.44
$K_{i,t-1}$	-0.059	-0.072	-0.059	-0.046	1.00	100.00
$I^N_{i,t-1}$	-0.055	-0.094	-0.057	0.000	1.00	94.00
$\Delta Q_{i,t}^2$	0.000	0.000	0.000	0.000	1.00	0.57
$V_{i,t}^2$	0.000	0.000	0.000	0.000	1.00	1.93
$C_{i,t}^2$	-0.003	-0.030	0.000	0.000	1.00	18.37
$AGE_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.23
$TL_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.06
$NW_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.03
$CL_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.75
$K_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.51
$I^N{}_{i,t-1}^2$	0.020	0.014	0.021	0.026	1.01	99.14
$\Delta Q_{i,t}*V_{i,t}$	0.000	0.000	0.000	0.000	1.00	0.58
$\Delta Q_{i,t}*C_{i,t}$	0.000	0.000	0.000	0.000	1.00	2.64
$\Delta Q_{i,t}*AGE_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.47
$\Delta Q_{i,t}*TL_{i,t-1}$	0.000	-0.001	0.000	0.000	1.00	45.97
$\Delta Q_{i,t}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.24
$\Delta Q_{i,t}*CL_{i,t-1}$	0.000	-0.002	0.000	0.000	1.00	29.93
$\Delta Q_{i,t}*K_{i,t-1}$	0.000	-0.001	0.000	0.000	1.02	3.34
$\Delta Q_{i,t}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	2.18
$V_{i,t}*C_{i,t}$	0.000	0.000	0.000	0.000	1.05	2.14
$V_{i,t}*AGE_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.30
$V_{i,t}*TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.26
$V_{i,t}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	2.65
$V_{i,t}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	1.96
$V_{i,t}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.50
$V_{i,t}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.01	1.73
$C_{i,t}*AGE_{i,t-1}$	0.000	0.000	0.000	0.000	1.01	1.38
$C_{i,t}*TL_{i,t-1}$	0.000	-0.003	0.000	0.000	1.00	10.30
$C_{i,t}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.63
$C_{i,t}*CL_{i,t-1}$	0.000	-0.005	0.000	0.000	1.00	6.44

Table A2.8. (continued)

Variable	Posterior Mean	Posterior quantiles 2.5%	50%	97.5%	$\sqrt{\hat{R}}$	% of times chosen
$C_{i,t}*K_{i,t-1}$	-0.003	-0.011	0.000	0.000	1.00	36.05
$C_{i,t}*I^N_{i,t-1}$	-0.001	-0.016	0.000	0.000	1.00	11.62
$AGE_{i,t-1}*TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.30
$AGE_{i,t-1}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.10
$AGE_{i,t-1}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.58
$AGE_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.94
$AGE_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.002	1.00	4.55
$TL_{i,t-1}*NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.36
$TL_{i,t-1}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.20
$TL_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.30
$TL_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.07	1.80
$NW_{i,t-1}*CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.82
$NW_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.10
$NW_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.51
$CL_{i,t-1}*K_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.60
$CL_{i,t-1}*I^N_{i,t-1}$	0.001	0.000	0.000	0.007	1.00	11.93
$K_{i,t-1}*I^N_{i,t-1}$	0.000	0.000	0.000	0.000	1.07	2.93
Parameter						
η	0.710	0.671	0.710	0.747	1.00	
σ_ε^2	0.863	0.770	0.861	0.968	1.00	
σ_v^2	0.476	0.253	0.445	0.892	1.01	

APPENDIX 3. SUBMODEL PARAMETER ESTIMATES

Table A3.1. Submodel parameter estimates

Parameter	Value	S. E.	Value	S. E.	Value	S. E.
Adj. costs						
θ_0	0.009	0.054	0.138^{*}	0.062	0.049	0.060
Fin. constraint						
ρ_0	-0.002	0.033	0.042	0.022	0.046	0.033
ρ_1 (AGE_{t-1})			-0.0033	0.0034	-0.0016	0.0026
ρ_2 (V_t)	0.0091^{*}	0.0040			0.0041	0.0026
ρ_3 (TL_{t-1})	0.0007	0.0031	0.0051	0.0027		
ρ_4 (NW_{t-1})	-0.0028	0.0017	-0.0007	0.0013	-0.0021	0.0018
ρ_5 (CL_{t-1})	0.0090	0.0088	-0.0083	0.0068	0.0116	0.0062
ρ_6 $(AGE_{t-1})^2$			$-2.5\text{x}10^{-4}$	$1.8\text{x}10^{-4}$	$-1.4\text{x}10^{-4}$	$1.5\text{x}10^{-4}$
ρ_7 $(V_t)^2$	$-3.0\text{x}10^{-4*}$	$1.2\text{x}10^{-4}$			$-1.82\text{x}10^{-4*}$	$7.9\text{x}10^{-5}$
ρ_8 $(TL_{t-1})^2$	$9.0\text{x}10^{-5}$	$7.5\text{x}10^{-5}$	$-6.3\text{x}10^{-5}$	$4.5\text{x}10^{-5}$		
ρ_9 $(NW_{t-1})^2$	$9.9\text{x}10^{-5*}$	$4.0\text{x}10^{-5}$	$6.5\text{x}10^{-6}$	$9.4\text{x}10^{-6}$	$7.5\text{x}10^{-4*}$	$3.3\text{x}10^{-5}$
ρ_{10} $(CL_{t-1})^2$	$-2.09\text{x}10^{-3*}$	$8.9\text{x}10^{-4}$	$3.2\text{x}10^{-4}$	$3.4\text{x}10^{-4}$	$-1.70\text{x}10^{-3*}$	$7.3\text{x}10^{-4}$

J-stat. tests[a]		p-value		p-value		p-value
χ^2 (O. R.)	6.56	0.5846	9.93	0.2702	7.91	0.4419
d. f.	8		8		8	
$\bar{\chi}^2$ (vs. full)	1.84	0.3991	5.20	0.0742	3.19	0.2029
d. f.	2		2		2	

[a]O. R. stands for overidentifying restrictions, full stands for Model 3 in Table 6.2. The parameter estimates for the time effects are not presented.

*Significantly different from zero at the 5% level based on the two-tailed *t*-statistic.

Table A3.1. (continued)

Parameter	Value	S. E.	Value	S. E.	Value	S. E.
Adj. costs						
θ_0	0.129*	0.058	0.141*	0.063	0.122	0.048
Fin. constraint						
ρ_0	0.035	0.039	0.038	0.044	0.039	0.014
ρ_1 (AGE_{t-1})	-0.0027	0.0031	-0.0066	0.0036		
ρ_2 (V_t)	0.0000	0.0022	0.0009	0.0028		
ρ_3 (TL_{t-1})	0.0039	0.0031	0.0013	0.0025	0.0048	0.0025
ρ_4 (NW_{t-1})			-0.0011	0.0020	-0.00089	0.00090
ρ_5 (CL_{t-1})	-0.0084	0.0068			-0.0070	0.0064
ρ_6 $(AGE_{t-1})^2$	$-2.1\text{x}10^{-4}$	$1.7\text{x}10^{-4}$	$-3.8\text{x}10^{-4}$*	$1.9\text{x}10^{-4}$		
ρ_7 $(V_t)^2$	$0.4\text{x}10^{-5}$	$3.6\text{x}10^{-5}$	$-4.0\text{x}10^{-5}$	$5.2\text{x}10^{-5}$		
ρ_8 $(TL_{t-1})^2$	$2.4\text{x}10^{-5}$	$7.2\text{x}10^{-5}$	$0.4\text{x}10^{-5}$	$7.9\text{x}10^{-5}$	$-7.5\text{x}10^{-5}$	$3.7\text{x}10^{-5}$
ρ_9 $(NW_{t-1})^2$			$1.5\text{x}10^{-4}$	$1.6\text{x}10^{-5}$	$1.0\text{x}10^{-5}$	$7.3\text{x}10^{-6}$
ρ_{10} $(CL_{t-1})^2$	$3.6\text{x}10^{-4}$	$2.5\text{x}10^{-4}$			$1.2\text{x}10^{-4}$	$2.6\text{x}10^{-4}$

J-stat. tests[a]		p-value		p-value		p-value
χ^2 (O. R.)	10.88	0.2086	8.922	0.3489	15.49	0.1153
d. f.	8		8		10	
χ^2 (vs. full)	6.15	0.0461	4.20	0.1226	10.76	0.0294
d. f.	2		2		4	

[a]O. R. stands for overidentifying restrictions, full stands for Model 3 in Table 6.2. The parameter estimates for the time effects are not presented.

*Significantly different from zero at the 5% level based on the two-tailed *t*-statistic.

Table A3.1. (continued)

Parameter	Value	S. E.	Value	S. E.	Value	S. E.
Adj. costs						
θ_0	0.015	0.052	0.113*	0.047	0.117*	0.048
Fin. constraint						
ρ_0	0.035	0.025	0.021	0.031	0.040	0.033
ρ_1 (AGE_{t-1})						
ρ_2 (V_t)	0.0046*	0.0023	0.0005	0.0019	-0.0001	0.0022
ρ_3 (TL_{t-1})			0.0035	0.0028	0.0000	0.0022
ρ_4 (NW_{t-1})	-0.0019	0.0015			-0.0005	0.0012
ρ_5 (CL_{t-1})	0.0107	0.0059	-0.0087	0.0067		
ρ_6 $(AGE_{t-1})^2$						
ρ_7 $(V_t)^2$	$-1.98\text{x}10^{-4}$*	$7.9\text{x}10^{-5}$	$-0.9\text{x}10^{-5}$	$3.2\text{x}10^{-5}$	$-0.7\text{x}10^{-5}$	$3.2\text{x}10^{-5}$
ρ_8 $(TL_{t-1})^2$			$-1.7\text{x}10^{-5}$	$6.2\text{x}10^{-5}$	$1.3\text{x}10^{-5}$	$6.7\text{x}10^{-5}$
ρ_9 $(NW_{t-1})^2$	$8.0\text{x}10^{-5}$*	$3.3\text{x}10^{-5}$			$1.8\text{x}10^{-6}$	$5.9\text{x}10^{-6}$
ρ_{10} $(CL_{t-1})^2$	$-1.80\text{x}10^{-3}$*	$7.1\text{x}10^{-4}$	$2.8\text{x}10^{-4}$	$2.4\text{x}10^{-4}$		

J-stat. tests[a]		p-value		p-value		p-value
χ^2 (O. R.)	8.53	0.5774	15.49	0.1152	15.99	0.0998
d. f.	10		10		10	
χ^2 (vs. full)	3.80	0.4332	10.77	0.0293	11.27	0.0237
d. f.	4		4		4	

[a]O. R. stands for overidentifying restrictions, full stands for Model 3 in Table 6.2. The parameter estimates for the time effects are not presented.

*Significantly different from zero at the 5% level based on the two-tailed *t*-statistic.

Table A3.1. (continued)

Parameter	Value	S. E.	Value	S. E.	Value	S. E.
Adj. costs						
θ_0	0.115*	0.054	0.137*	0.047	0.127*	0.051
Fin. constraint						
ρ_0	0.033	0.020	0.037*	0.016	0.034	0.021
ρ_1 (AGE_{t-1})	-0.0035	0.0034	-0.0024	0.0028	-0.0042	0.0025
ρ_2 (V_t)						
ρ_3 (TL_{t-1})			0.0042	0.0026	0.0026	0.0017
ρ_4 (NW_{t-1})	-0.0000	0.0011			-0.0006	0.0013
ρ_5 (CL_{t-1})	0.0001	0.0034	-0.0089	0.0064		
ρ_6 $(AGE_{t-1})^2$	$-2.0\text{x}10^{-4}$	$1.7\text{x}10^{-4}$	$-2.0\text{x}10^{-4}$	$1.5\text{x}10^{-4}$	$-2.5\text{x}10^{-4}$	$1.3\text{x}10^{-4}$
ρ_7 $(V_t)^2$						
ρ_8 $(TL_{t-1})^2$			$-3.5\text{x}10^{-5}$	$2.0\text{x}10^{-5}$	$-4.8\text{x}10^{-5}$	$3.9\text{x}10^{-5}$
ρ_9 $(NW_{t-1})^2$	$0.3\text{x}10^{-6}$	$3.6\text{x}10^{-6}$			$7.1\text{x}10^{-6}$	$9.6\text{x}10^{-6}$
ρ_{10} $(CL_{t-1})^2$	$0.9\text{x}10^{-4}$	$2.0\text{x}10^{-4}$	$3.7\text{x}10^{-4}$	$2.4\text{x}10^{-4}$		

J-stat. tests[a]		p-value		p-value		p-value
χ^2 (O. R.)	13.28	0.2083	11.60	0.3124	10.85	0.3696
d. f.	10		10		10	
χ^2 (vs. full)	8.56	0.0732	6.88	0.1424	6.12	0.1903
d. f.	4		4		4	

[a]O. R. stands for overidentifying restrictions, full stands for Model 3 in Table 6.2. The parameter estimates for the time effects are not presented.
*Significantly different from zero at the 5% level based on the two-tailed *t*-statistic.

Table A3.1. (continued)

Parameter	Value	S. E.	Value	S. E.	Value	S. E.
Adj. costs						
θ_0	0.108	0.058	-1.89*	0.74	0.111*	0.042
Fin. constraint						
ρ_0	0.063	0.040	-13	31	0.034*	0.013
ρ_1 (AGE_{t-1})	-0.0038	0.0028	-0.03	0.15		
ρ_2 (V_t)	-0.0012	0.0020	5	10		
ρ_3 (TL_{t-1})			0.14	0.28		
ρ_4 (NW_{t-1})	0.0000	0.0016			-0.00042	0.00066
ρ_5 (CL_{t-1})					-0.0003	0.0030
ρ_6 $(AGE_{t-1})^2$	-0.00023	0.00016	-0.010	0.016		
ρ_7 $(V_t)^2$	$-0.9\text{x}10^{-5}$	$2.3\text{x}10^{-5}$	-0.37	0.82		
ρ_8 $(TL_{t-1})^2$			-0.007	0.016		
ρ_9 $(NW_{t-1})^2$	$0.6\text{x}10^{-5}$	$1.2\text{x}10^{-5}$			$1.1\text{x}10^{-6}$	$2.2\text{x}10^{-6}$
ρ_{10} $(CL_{t-1})^2$					0.00002	0.00014

J-stat. tests[a]		p-value		p-value		p-value
χ^2 (O. R.)	11.49	0.3203	7.72	0.6559	19.03	0.0877
d. f.	10		10		12	
χ^2 (vs. full)	6.77	0.1486	3.00	0.5582	14.31	0.0264
d. f.	4		4		6	

[a]O. R. stands for overidentifying restrictions, full stands for Model 3 in Table 6.2. The parameter estimates for the time effects are not presented.

*Significantly different from zero at the 5% level based on the two-tailed *t*-statistic.

Table A3.1. (continued)

Parameter	Value	S. E.	Value	S. E.	Value	S. E.
Adj. costs						
θ_0	0.114*	0.044	0.132*	0.046	0.109*	0.045
Fin. constraint						
ρ_0	0.0284*	0.0074	0.038*	0.014	0.050	0.029
ρ_1 (AGE_{t-1})						
ρ_2 (V_t)					-0.0010	0.0017
ρ_3 (TL_{t-1})	0.0037	0.0024	0.00082	0.00091		
ρ_4 (NW_{t-1})			-0.00068	0.00081	-0.0002	0.0011
ρ_5 (CL_{t-1})	-0.0082	0.0065				
ρ_6 $(AGE_{t-1})^2$						
ρ_7 $(V_t)^2$					$3.4\text{x}10^{-5}$	$1.2\text{x}10^{-5}$
ρ_8 $(TL_{t-1})^2$	$-3.1\text{x}10^{-5}$	$1.9\text{x}10^{-5}$	$-0.8\text{x}10^{-5}$	$1.2\text{x}10^{-5}$		
ρ_9 $(NW_{t-1})^2$			$2.3\text{x}10^{-6}$	$4.9\text{x}10^{-6}$	$1.2\text{x}10^{-6}$	$5.7\text{x}10^{-6}$
ρ_{10} $(CL_{t-1})^2$	0.00027	0.00024				

J-stat. tests[a]		p-value		p-value		p-value
χ^2 (O. R.)	15.49	0.2156	17.67	0.1261	16.17	0.1836
d. f.	12		12		12	
χ^2 (vs. full)	10.77	0.0958	12.94	0.0439	11.44	0.0756
d. f.	6		6		6	

[a]O. R. stands for overidentifying restrictions, full stands for Model 3 in Table 6.2. The parameter estimates for the time effects are not presented.
*Significantly different from zero at the 5% level based on the two-tailed *t*-statistic.

Table A3.1. (continued)

Parameter	Value	S. E.	Value	S. E.	Value	S. E.
Adj. costs						
θ_0	0.109*	0.042	0.139*	0.047	0.107*	0.037
Fin. constraint						
ρ_0	0.033*	0.014	0.032*	0.014	0.032*	0.011
ρ_1 (AGE_{t-1})	-0.0007	0.0010	-0.0016	0.0015		
ρ_2 (V_t)						
ρ_3 (TL_{t-1})			0.0014	0.0013		
ρ_4 (NW_{t-1})	0.00006	0.00049			-0.00025	0.00035
ρ_5 (CL_{t-1})						
ρ_6 $(AGE_{t-1})^2$	$-4.5\text{x}10^{-5}$	$8.3\text{x}10^{-5}$	$-1.25\text{x}10^{-4}$	$9.6\text{x}10^{-5}$		
ρ_7 $(V_t)^2$						
ρ_8 $(TL_{t-1})^2$			$-1.2\text{x}10^{-5}$	$1.1\text{x}10^{-5}$		
ρ_9 $(NW_{t-1})^2$	$-0.5\text{x}10^{-6}$	$1.7\text{x}10^{-6}$			$0.6\text{x}10^{-6}$	$1.2\text{x}10^{-6}$
ρ_{10} $(CL_{t-1})^2$						

J-stat. tests[a]		p-value		p-value		p-value
χ^2 (O. R.)	18.17	0.1107	14.72	0.2570	18.87	0.1701
d. f.	12		12		14	
χ^2 (vs. full)	13.44	0.0365	10.00	0.1248	14.14	0.0781
d. f.	6		6		8	

[a]O. R. stands for overidentifying restrictions, full stands for Model 3 in Table 6.2. The parameter estimates for the time effects are not presented.

*Significantly different from zero at the 5% level based on the two-tailed *t*-statistic.

Table A3.1. (continued)

Parameter	Value	S. E.	Value	S. E.
Adj. costs				
θ_0	0.100*	0.040	0.070*	0.031
Fin. constraint				
ρ_0	0.0253*	0.0077	0.024*	0.012
ρ_1 (AGE_{t-1})			0.00025	0.00069
ρ_2 (V_t)				
ρ_3 (TL_{t-1})	0.00021	0.00078		
ρ_4 (NW_{t-1})				
ρ_5 (CL_{t-1})				
ρ_6 $(AGE_{t-1})^2$			$-8.7\text{x}10^{-5}$	$5.8\text{x}10^{-5}$
ρ_7 $(V_t)^2$				
ρ_8 $(TL_{t-1})^2$	$-2.8\text{x}10^{-6}$	$6.2\text{x}10^{-6}$		
ρ_9 $(NW_{t-1})^2$				
ρ_{10} $(CL_{t-1})^2$				

J-stat. tests[a]		p-value		p-value
χ^2 (O. R.)	18.80	0.1726	22.67	0.0659
d. f.	14		14	
χ^2 (vs. full)	14.08	0.0797	17.94	0.0217
d. f.	8		8	

[a]O. R. stands for overidentifying restrictions, full stands for Model 3 in Table 6.2. The parameter estimates for the time effects are not presented.
*Significantly different from zero at the 5% level based on the two-tailed t-statistic.

REFERENCES

Akaike, H. 1973. "Information Theory and an Extension of the Maximum Likelihood Principle." In 2^{nd} *International Symposium on Information Theory*, edited by B. N. Petrow and F. Csaki, pages 267-281. Budapest: Akademiai Kiodo.

Barnett, V. and T. Lewis. 1984. *Outliers in Statistical Data*. New York: Wiley.

Barran F. and M. Peeters. 1998. "Internal Finance and Corporate Investment: Belgian Evidence with Panel Data." *Economic Modelling* 15: 67-89.

Beckman, R. J. and R. D. Cook. 1983. "Outlier..........s." *Technometrics* 25: 119-149.

Berger, J. O. 1985. *Statistical Decision Theory and Bayesian Analysis*. New York: Springer-Verlag.

Bernardo, J. M. and A. F. M. Smith. 1994. *Bayesian Theory*. New York: Wiley.

Besag, J. 1974. "Spatial Interaction and the Statistical Analysis of Lattice Systems." *Journal of the Royal Statistical Society, Series B (Methodology)* 36: 192-225.

Bierlen, R. W. 1994. "The Interdependence of Investment and Financing Decisions among Kansas Farm Machinery Investors." Manuscript, Kansas State University, Manhattan.

Bierlen, R. W. and A. M. Featherstone. 1996. "Credit Rationing and Farm Machinery Investment." Manuscript, University of Arkansas, Fayetteville.

Bierlen, R. W. and A. M. Featherstone. 1998. "Fundamental q, Cash Flow, and Investment: Evidence from Farm Panel Data." *Review of Economics and Statistics* 80: 427-435.

Bond, S. and C. Meghir. 1994. "Dynamic Investment Models and the Firm's Financial Policy." *The Review of Economic Studies* 61: 197-222.

Box, G. E. P. and G. C. Tiao. 1968. "A Bayesian Approach to some Outlier Problems." *Biometrika* 55: 119-129

Box, G. E. P. and G. C. Tiao. 1973. *Bayesian Inference in Statistical Analysis*. Reading, MA: Addison-Wesley.

Brainard, W. C. and J. Tobin. 1968. "Pitfalls in Financial Model Building." *American Economic Review* 58: 99-122.

Brooks, S. P. 1998. "Markov Chain Monte Carlo Method and its Application." *The Statistician* 47: 69-100.

Carroll, C. D. 1997. "Death to the Log-Linearized Consumption Euler Equation! (And Very Poor Health to the Second-Order Approximation)." Working paper 6298. National Bureau of Economic Research, Inc.

Casella, G., and E. I. George. 1992. "Explaining the Gibbs Sampler." *The American Statistician* 46: 167-174.

Chellappan R. RM and G. D. Pederson. 1995. "Analysis of Investment Behavior on Minnesota Crop Farms." Selected paper presented at the Conference of the American Agricultural Economics Association at Indianapolis, Indiana, August 6-9.

Chenery, H. B. 1952. "Overcapacity and the Acceleration Principle." *Econometrica* 20: 1-28.

Chirinko, R. S. 1993. "Business Fixed Investment Spending: A Critical Survey of Modeling Strategies, Empirical Results, and Policy Implications." Working paper RWP 93-01. Federal Reserve Bank of Kansas City.

Chirinko, R. S. and H. Schaller. 1996. "Bubbles, Fundamentals, and Investment: A Multiple Equation Testing Strategy" *Journal of Monetary Economics* 38: 47-76.

Chirinko, R. S. and H. Schaller. 1995. "Why Does Liquidity Matter in Investment Equations?" *Journal of Money, Credit, and Banking* 27: 527-548.

Chow, G. C. 1981. "A Comparison of the Information and Posterior Probability Criteria for Model Selection." *Journal of Econometrics* 16: 21-33.

Chow, G. C. 1983. *Econometrics*. New York: McGraw-Hill.

Clark, J. M. 1917. "Business Acceleration and the Law of Demand: A Technical Factor in Economic Cycles." *Journal of Political Economy* 25: 217-235.

Davidson, R. and J. G. MacKinnon. 1993. *Estimation and Inference in Econometrics*. New York: Oxford University Press.

Estrada, Á. And J. Vallés. 1995. "Investment and Financial Costs: Spanish Evidence with Panel Data." Working paper 9506. Banco de España.

Everitt, B. S. and G. Dunn. 1991. *Applied Multivariate Data Analysis*. London: Edward Arnold.

Fazzari, S. M. and B. C. Petersen. 1993. "Working Capital and Fixed Investment: New Evidence on Financing Constraints." *RAND Journal of Economics* 24: 328-342.

Fazzari, S. M. and M. J. Athey. 1987. "Asymmetric Information, Financing Constraints, and Investment." *Review of Economics and Statistics* 69: 481-487.

Fazzari, S. M. and T. L. Mott. 1986. "The Investment Theories of Kalecki and Keynes: An Empirical Study of Firm Data." *Journal of Post-Keynesian Economics* 9: 171-187.

Fazzari, S. M., R. G. Hubbard, and B. C. Petersen. 1988. "Financing Constraints and Corporate Investment." *Brookings Papers on Economic Activity* 1: 141-195.

Feller, W. 1968. *An Introduction to Probability Theory and its Application.* 3rd edition. New York: John Wiley and Sons.

Fuhrer, J. C., G. R. Moore, and S. D. Schuh. 1995. "Estimating the Linear-Quadratic Inventory Model: Maximum Likelihood Versus Generalized Method of Moments." *Journal of Monetary Economics* 35: 115-157.

Gelfand, A. E. and A. F. M. Smith. 1990. "Sampling-Based Approaches to Calculating Marginal Densities." *Journal of the American Statistical Association* 85: 398-409.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 1995. *Bayesian Data Analysis.* New York: Chapman and Hall.

Geman, S. and D. Geman. 1984. "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6: 721-741.

George, E. I. and R. E. McCulloch. 1993. "Variable Selection Via Gibbs Sampling." *Journal of the American Statistical Association* 88: 881-889.

George, E. I. and R. E. McCulloch. 1997. "Approaches for Bayesian Variable Selection." *Statistica Sinica* 7: 339-373.

Geweke, J. K. 1996. "Variable Selection and Model Comparison in Regression." In *Bayesian Statistics 5*, edited by J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, pages 609-620. Oxford: Oxford Press.

Gilchrist, S. and C. P. Himmelberg. 1995. "Evidence on the Role of Cash Flow for Investment." *Journal of Monetary Economics* 36: 541-572.

Goolsbee, A. and D. B. Gross. 1997. "Estimating Adjustment Costs with Data on Heterogeneous Capital Goods." Working paper 6342. National Bureau of Economic Research, Inc.

Gourieroux, C. and A. Monfort. 1995. *Statistics and Econometric Models: Volume 2: Testing, Confidence Regions, Model Selection, and Asymptotic Theory*. Cambridge: Cambridge University Press.

Grasa, A. A. 1989. *Econometric Model Selection: A New Approach*. Dordrecht: Kluwer Academic Publishers.

Greene, W. H. 1990. *Econometric Analysis*. New York: Macmillan Publishing Company.

Gustafson, C. R., P. J. Barry, and S. T. Sonka. 1988. "Machinery Investment Decisions: A Simulated Analysis for Cash Grain Farms." *Western Journal of Agricultural Economics* 13: 244-253.

Hamermesh, D. S. and G. A. Pfann. 1996. "Adjustment Costs in Factor Demand." *Journal of Economic Literature* 34: 1264-1292.

Hansen, L. P. 1982. "Large Sample Properties of Generalized Method of Moments Estimators." *Econometrica* 50: 1029-1054.

Hansen, L. P. and K. J. Singleton. 1982. "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models." *Econometrica* 50: 1269-1286.

Hastings, W. K. 1970. "Monte Carlo Sampling Methods using Markov Chains and Their Applications." *Biometrika* 57: 97-109.

Hayashi, F. 1982. "Tobin's Marginal q and Average q: A Neoclassical Interpretation." *Econometrica* 50: 213-224.

Hayashi, F. 1985. "Corporate Finance Side of the Q Theory of Investment." *Journal of Public Economics* 27: 261-280.

Hoeting, J., A. E. Raftery, and D. Madigan. 1996. "A Method for Simultaneous Variable Selection and Outlier Identification in Linear Regression." *Computational Statistics and Data Analysis* 22: 251-270.

Hoeting, J., D. Madigan, A. E. Raftery, and C. T. Volinsky. 1998. "Bayesian Model Averaging." Technical Report 335. Department of Statistics. University of Washington, Seattle.

Hubbard, R. G. and A. K. Kashyap. 1992. "Internal Net Worth and the Investment Process: An Application to U.S. Agriculture." *Journal of Political Economy* 100: 506-534.

Hubbard, R. G., A. K. Kashyap, and T. M. Whited. 1995. "Internal Finance and Firm Investment." *Journal of Money, Credit, and Banking* 27: 683-701.

Hughes, J. P. and L. J. Mester. 1995. "Bank Capitalization and Cost: Evidence of Scale Economies in Risk Management and Signaling." Working paper 96-2. Federal Reserve Bank of Philadelphia.

Iglewicz, B. and D. C. Hoaglin. 1993. *How to Detect and Handle Outliers*. Milwaukee, WI: ASQC Quality Press.

Jensen, F. E., J. S. Lawson, and L. N. Langemeier. 1993. "Agricultural Investment and Internal Cash Flow Variables." *Review of Agricultural Economics* 15: 295-306.

Jorgenson, D. W. 1963. "Capital Theory and Investment Behavior." *American Economic Review* 53: 247-259.

Jorgenson, D. W. 1971. "Econometric Studies of Investment Behavior: A Survey." *Journal of Economic Literature* 9: 1111-1147.

Keynes, J. M. 1936. *The General Theory of Employment, Interest, and Money*. New York: Harcourt Brace.

Klein, L. R. 1951. "Studies in Investment Behavior." In *Conference on Business Cycles*, pages 233-303. New York: National Bureau of Economic Research.

Knopf, E. and R. Schoney. 1993. "An Evaluation of Farm Financial Benchmarks and Loan Success/Failure: The Case of the Agricultural Credit Corporation of Saskatchewan." *Canadian Journal of Agricultural Economics* 41: 61-69.

Koyck, L. M. 1954. *Distributed Lags and Investment Analysis*. Amsterdam: North-Holland.

LaDue, E. L., L. H. Miller, and J. H. Kwiatkowski. 1991. "Investment Behavior and Farm Business Expansion." *Review of Agricultural Economics* 13: 73-84.

Lahiri, K. 1993. "Panel Data Models with Rational Expectations." In *Handbook of Statistics, Volume 11: Econometrics*, edited by G. S. Maddala, C. R. Rao, and H. D. Vinod, Chapter 26. New York: North-Holland.

Madigan, D. and A. E. Raftery. 1994. "Model Selection and Accounting for Model Uncertainty in Graphical Models Using Occam's Window." *Journal of the American Statistical Association* 89: 1535-1546.

Metropolis, N. and S. Ulam. 1949. "The Monte Carlo Method." *Journal of the American Statistical Association* 44: 335-341.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. 1953. "Equation of State Calculations by Fast Computing Machines." *Journal of Chemical Physics* 21: 1087-1092.

Miller, A. J. 1990. *Subset Selection in Regression.* New York: Chapman and Hall.

Miller, L. H. and E. L. LaDue. 1989. "Credit Assessment Models for Farm Borrowers: A Logit Analysis." *Agricultural Finance Review* 49: 22-36.

Miller, M. H. and F. Modigliani. 1961. "Dividend Policy, Growth, and the Valuation of Shares." *Journal of Business* 34: 411-433.

Modigliani, F. and M. H. Miller. 1963. "Corporate Income Taxes and the Cost of Capital: A Correction." *American Economic Review* 53: 433-443.

Modigliani, F. and M. H. Miller. 1958. "The Cost of Capital, Corporation Finance and the Theory of Investment." *American Economic Review* 48: 261-297.

Nelson, C. R. and R. Startz. 1990. "The Distribution of the Instrumental Variables Estimator and Its *t*-Ratio When the Instrument is a Poor One." *Journal of Business* 63: S125-S140.

Ogaki, M. 1993. "Generalized Method of Moments: Econometric Applications." In *Handbook of Statistics, Volume 11: Econometrics*, edited by G. S. Maddala, C. R. Rao, and H. D. Vinod, Chapter 17. New York: North-Holland.

Oliner, S. D., G. D. Rudebusch, and D. Sichel. 1995. "New and Old Models of Business Investment: A Comparison of Forecasting Performance." *Journal of Money, Credit, and Banking* 27: 806-826.

Oliner, S. D., G. D. Rudebusch, and D. Sichel. 1996. "The Lucas Critique Revisited: Assessing the Stability of Empirical Euler Equations for Investment." *Journal of Econometrics* 70: 291-316.

Raftery, A. E., D. Madigan, and C. T. Volinsky. 1996. "Accounting for Model Uncertainty in Survival Analysis Improves Predictive Performance." In *Bayesian Statistics 5*, edited by J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, pages 323-349. Oxford: Oxford Press.

Raftery, A. E., D. Madigan, and J. Hoeting. 1997. "Bayesian Model Averaging for Linear Regression Models." *Journal of the American Statistical Association* 92: 179-191.

Salinger, M. A. and L. H. Summers. 1983. "Tax Reform and Corporate Investment: A Microeconomic Simulation Study." In *Behavioral Simulation Methods in Tax Policy Analysis*, edited by M. Feldstein, Chapter 8. Chicago: University of Chicago Press.

Sawa, T. 1978. "Information Criteria for Discriminating Among Alternative Regression Models." *Econometrica* 46: 1273-1291.

Schwarz, G. 1978. "Estimating the Dimension of a Model." *The Annals of Statistics* 6: 461-464.

Snedecor, G. W. and W. G. Cochran. 1989. *Statistical Methods*. 8th edition. Ames: Iowa State University Press.

Stiglitz, J. E. 1969. "A Re-Examination of the Modigliani-Miller Theorem." *American Economic Review* 59: 784-793.

Tinbergen, J. 1938. "Statistical Evidence on the Acceleration Principle." *Econometrica* 5: 164-176.

Tinbergen, J. 1939. "A Method and its Application to Investment Activity." In *Statistical Testing of Business Cycle Theories, Volume 1*. Geneva: League of Nations.

Tobin, J. 1969. "A General Equilibrium Approach to Monetary Theory." *Journal of Money, Credit, and Banking* 1: 15-29.

U.S. Department of Agriculture, National Agricultural Statistics Service, Agricultural Statistics Board. *Agricultural Prices*. Various issues. Washington, D.C.

U.S. Department of Commerce, Economics and Statistics Administration, Bureau of Economic Analysis. *Survey of Current Business*. Washington, D.C.

U.S. Department of Commerce, Economics and Statistics Administration, Bureau of the Census. *1992 Census of Agriculture*. Washington, D.C.

Walraven, N. A. and D. Carson. 1996. *Agricultural Finance Databook.* Division of Research and Statistics, Board of Governors of the Federal Reserve System. Washington, D.C.

Weersink, A. J. and L. W. Tauer. 1989. "Comparative Analysis of Investment Models for New York Dairy Farms." *American Journal of Agricultural Economics* 71: 136-146.

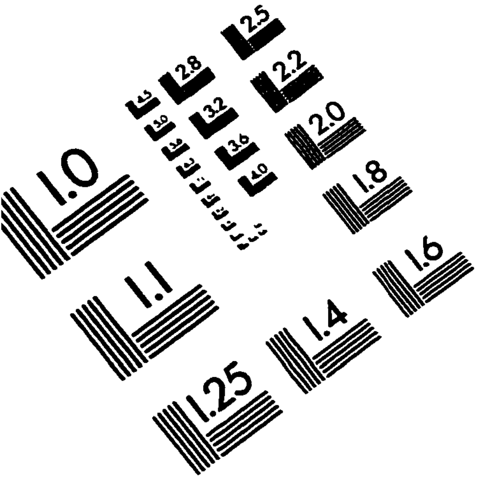
White, H. 1980. "A Heteroscedasticity-Consistent Covariance Matrix and a Direct Test for Heteroscedasticity." *Econometrica* 48: 817-838.

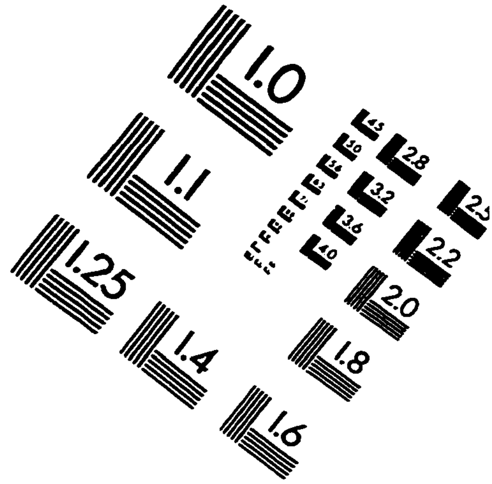
Whited, T. M. 1992. "Debt, Liquidity Constraints, and Corporate Investment: Evidence from Panel Data." *Journal of Finance* 47: 1425-1460.

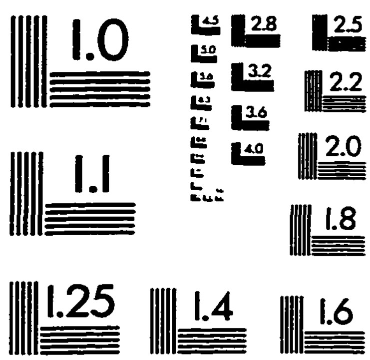
ACKNOWLEDGEMENTS

There are many people I would like to thank for their support during my time as a graduate student at Iowa State University. First, my most sincere appreciation goes to my co-major professors, Drs. Sergio H. Lence and Alicia L. Carriquiry. Their guidance through the successes and failures of my research and the process of obtaining co-majors in Agricultural Economics and Statistics has been invaluable. Next, I would like to thank my committee members, Drs. Wayne Fuller, Arne Hallam, and Elgin Johnston, for contributing their expertise and time to my work. My supervisors at CARD and FAPRI, Drs. Dermot Hayes, Darnell Smith, William Meyers, and Bruce Babcock, have been extremely supportive of my personal and professional work. My experience with CARD and FAPRI has allowed me to explore many issues in agricultural economics in a short amount of time. Finally, I would like to thank Jennifer, my wife, for always being here for me. I would not have survived this without her.

IMAGE EVALUATION TEST TARGET (QA-3)

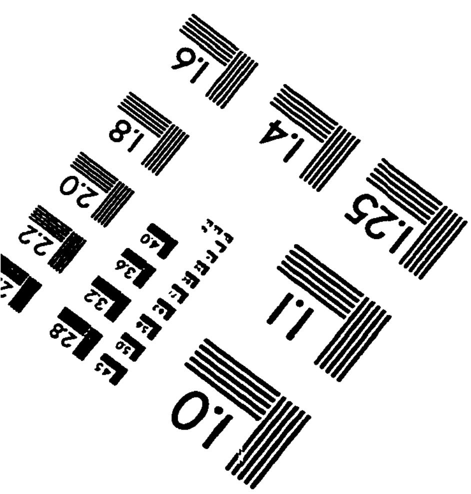


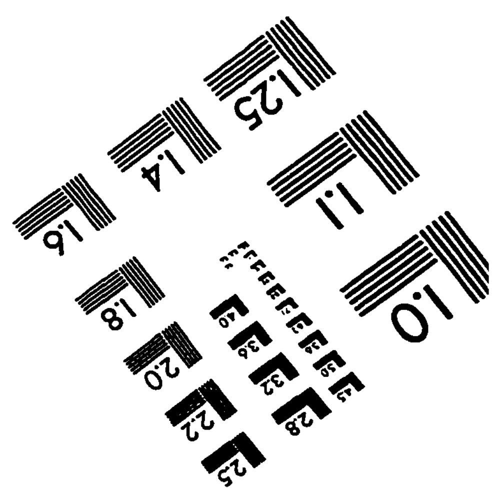




150mm

6"





APPLIED IMAGE . Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989