THE DEMAND FOR MONEY BY FIRMS

By

JEFFREY M. HERBENER

Bachelor of Science
Nebraska Wesleyan University
Lincoln, Nebraska
1977

Master of Science
Oklahoma State University
Stillwater, Oklahoma
1981

Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
DOCTOR OF PHILOSOPHY
December, 1982

THE DEMAND FOR MONEY BY FIRMS

Thesis Approved:

John D. Rea

Thesis Adviser

Frank G. Steindl

Michael R. Edgmand

Jim C. Freeland

Norman N. Durham

Dean of the Graduate College

ACKNOWLEDGMENTS

I must extend appreciation to my major adviser, Dr. John Rea, for invaluable assistance, direction, and motivation. Initial admiration for his academic skills has deepened continuously through the course of this work. He has awakened me and started a growth I hope will never end. I wish someday to earn the right to call him a colleague. Thank you, Dr. Rea.

I hold the same wish with respect to the remaining committee members, Dr. Frank Steindl, Dr. Michael Edgmand, and Dr. Tim Ireland. Their selfless giving of time and extensive pool of knowledge coupled to render many crucial insights. I have been truly privileged to work with this committee and the faculty at Oklahoma State University.

Gratitude also goes out to Iris McPherson for generously sharing seemingly endless hours of instruction in implementing computer jobs. Marilynn Bond's tireless and always cheerful effort at the word processor deserves many thanks.

Special appreciation exists for my entire family, especially my father, Lloyd, and my mother, Ann, whose ceaseless encouragement and support provided the energy to endure. I cherish each of you.

TABLE OF CONTENTS

Chapter		Page
I.	INTRODUCTION	1
II.	REVIEW OF LITERATURE	5
	Introduction	5
	Inventory Model	6
	Theoretical Work	6
	Empirical Work	19
	Portfolio Model	23
	Theoretical Work	23
	Empirical Work	26
	Production Model	30
	Theoretical Work	30
	Empirical Work	36
	Summary	42
III.	DEVELOPMENT OF THE METHODOLOGY	43
	Introduction	43
	General Developments	46
	Overview	46
	Money	46
	Opportunity Cost of Money	47
	Summary	51
	Testable Specifications	52
	Inventory Model	52
	Portfolio Model	62
	Production Model	66
	Summary	70
	Econometric Techniques	71
	Overview	71
	Pooling Adjusted Methods	71
	Selecting a Specific Form	75
	Systems Methods	79
	Summary	81
	Samples	81
	Overview	81
	Quarterly Sample	82
	Annual Sample	83
	Summary	84
	Summary	89

Chapter Page

IV. ECONOMETRIC ANALYSIS 91

Introduction 91
Quarterly Sample 91
Annual Sample 111
Summary 131

V. SUMMARY AND CONCLUSIONS 134

A SELECTED BIBLIOGRAPHY 142

APPENDICES 144

APPENDIX A - OPPORTUNITY COST OF MONEY 153

APPENDIX B - EXTENSIONS OF THE INVENTORY MODEL 159

APPENDIX C - AGGREGATION ANALYSIS 180

APPENDIX D - DATA 191

APPENDIX E - TRANSLOG COST MODEL 195

APPENDIX F - ESTIMATION OF THE TRANSLOG COST MODEL 213

APPENDIX G - TEST PROCEDURES 221

LIST OF TABLES

Table		Page
I.	Variables	92
II.	Ordinary Least Squares (Quarterly)	94
III.	Generalized Difference Model (Quarterly)	97
IV.	Pooling Adjusted Models (Quarterly)	99
V.	Prediction Results (Quarterly)	106
VI.	Translog Cost Model	115
VII.	Ordinary Least Squares (Annual)	118
VIII.	Generalized Difference Model (Annual)	119
IX.	Pooling Adjusted Models (Annual)	122
X.	Prediction Results (Annual)	128

LIST OF FIGURES

Figure		Page
1.	Average Production Costs	167
2.	Average Transactions Costs	168
3.	Average Total Costs	169

LIST OF SYMBOLS

AD - Accumulated Depreciation

AP - Accounts Payable

AR - Accounts Receivable

ATC - Average Transfer Costs

ATRC - Average Transactions Costs

C_i - Relative Cost Share of the i^{th} Factor Input

CA - Real Current Assets

CBR - Ten Year Corporate Bond Rate

CE - Real Cash Expenditures

CF - Real Cash Flows

CMT - Cash Management Techniques

COGS - Cost of Goods Sold

COVM - Covariance Model

CR - Real Cash Receipts

D - Depreciation

ECM - Error Components Model

EQR - Equity Returns for Individual Firms

FFR - Federal Funds Rate

GDM - Generalized Difference Model

GI - Gross Investment

GPPE - Gross Property Plant and Equipment

I - Inventories

M - Real Money Balances

NI - Net Income

OLS - Ordinary Least Squares

PPC - Product Price Index for Specific Commodities

PPI - Product Price Index for Specific Industries

Q - Physical Output

RP - Repurchase Agreement

RPV - Inverse of Real Repurchase Agreement Volume

RR - Price of Capital

S - Real Sales

T - Income Taxes

V - Real Cash Flow Variance

WPI - Wholesale Price Index

WR - Wage Rate for Specific Commodities

CHAPTER I

INTRODUCTION

The purpose of this dissertation is to determine the most useful theory of a firm's money demand. The significance of making the choice stems from the two functions of any theory. First, a useful model correctly explains economic activity which deepens the understanding of that corner of the economic environment. In turn, the added knowledge sharpens perceptions of how the piece relates to the whole through a chain of cause and effect. Armed with new insights, better solutions can be found to deal with economic problems. Second, a useful theory makes accurate predictions of future or potential outcomes in economic activity. The quantity and price of money, through their effect on aggregate demand, play pervasive roles in determining macroeconomic activity and, therefore, society's welfare. To promote fully the well-being of individuals, the monetary authorities must do more than effectively control the money supply. They must also accurately predict money demand responses to changing economic circumstances since supply and demand jointly determine the quantity and price of money. Because firms hold a large portion of total private sector money balances, a model accurately predicting their behavior is an indispensable input to policy.

The recent case of the missing money highlights the importance of employing the correct model. As outlined by Goldfeld (1976), money demand functions, reliable predictors prior to 1974, began overestimating actual outcomes in the post-1974 period. Porter and Mauskopf (1979) along with Tinsley and Garrett (1978) attributed the error to changing behavior of firms in managing cash holdings. A monetary authority, unaware of the structural change, could establish money supply growth targets inconsistent with policy goals. Such an error would result in harmful variations in macroeconomic activity. In light of this crucial role, selecting the most useful description of firm's money demand is a paramount current issue.

The three competing theories, inventory, portfolio, and production, originated with Baumol (1952), Friedman (1956), and Gabor and Pierce (1958), respectively. Each evolves from a microfoundation of rational behavior and, thus, has equal theoretical justification. In such a circumstance, selection of the most useful theory must rely on econometric analysis. The volumes of empirical studies on money demand suggest that this is not a novel idea. Yet, 25 years of trying has failed to establish a consensus not only on the most useful theory but about many of the major issues pertaining to money demand functions. These include such basic topics as: a definition of money, the opportunity cost of money, the distinction between real and nominal balances, and the nature of money's alternative assets.

Much of the disagreement found in empirical works about money demand stems from the inconsistent interpretations of these topics. In addition, empirical studies have commonly conducted tests with a specification other than the correct one developed by the corresponding theoretical literature. In conjunction these two faults generates an inadequate framework for model evaluation. This problem constitutes only one out of three sources causing inconclusiveness of empirical work.

Inadequate data have also contributed to the confusion. In most cases, aggregate observations form the basis for evaluating theories about individual behavior. Thus, unwanted empirical results refuting a theory could be blamed on using data incapable of truly reflecting individual firm behavior.

Finally, money demand tests have not uniformly employed the most powerful econometric techniques available. In other cases, correct techniques have been applied in incorrect ways. A synthesis on econometric methods would, in combination with a comparable specifications and adequate data, provide for accurate evaluations of money demand models.

To accomplish its purpose, the dissertation must establish a methodology based on these three areas. First, taking a consistent approach to the topics pertaining to money demand renders specifications capable of making useful model evaluations. Second, samples of data must contain individual firm observations to test the microfoundation theories. Finally, the most powerful econometric techniques should be applied to data in order to provide estimation and prediction statistics that reflect each model's performance. These statistics will be judged by both explanation and prediction criteria since they constitute the two functions of a useful theory.

The dissertation performs its major purpose using the following organization. Chapter II presents both theoretical and empirical work previously done about firm's money demand. For each theory, the studies culminate in an equation embodying the unique aspects of that theory.

Providing these specifications represents the theoretical work's major input to the dissertation. The empirical literature makes a contribution by highlighting the areas that the dissertation can improve upon. Chapter III is devoted to establishing the three part methodology. Chapter IV applies the methodology to evaluate the three theories as explanations of and predictors for money demand. This testing results in the achievement of the dissertation's purpose. Finally, Chapter V summarizes the dissertation and speculates on the direction of future work.

CHAPTER II

REVIEW OF LITERATURE

Introduction

The purpose of this chapter is to outline the background needed to provide useful comparisons of the three money demand theories. Such a foundation relies, in different ways, on existing work of both a theoretical and empirical nature.

The theoretical studies' importance lies in the description of the evolution of testable functions. Beginning with the original idea, subsequent improvements eventually create a synthesis embodying the unique aspects of the underlying theory in a testable form. The culmination of this work provides the basic specifications examined in the dissertation.

The empirical literature has lagged behind theoretical development. Thus, its main contribution rests with highlighting the areas that the dissertation can improve upon. Overcoming these faults develops a methodology in which useful model comparisons can occur.

The review of literature is organized as follows. The three money demand theories are outlined in chronological order of their origins. For each model, the major theoretical work appears first, followed by a summary section describing its implications on the dissertation. Next, the major empirical work is described and then summarized with emphasis on the faults to be corrected by the dissertation. After outlining the work on all three models, a summary section recaps the main points of the literature.

Inventory Model

Theoretical Work

Overview. Money holdings can be viewed as a problem in inventory management. Firms cannot guarantee that the inflow of receipts from selling goods will exactly coincide with the expenditure outflows made to produce the goods. Holding a stock of cash becomes rational in order to synchronize these flows. Acquiring this benefit imposes two separate costs on the holder. Firms give up interest income by not holding securities and a transfer fee exists when shifting wealth between money and bonds. The optimum inventory of cash occurs at the level which minimizes the sum of these costs.

Deterministic Variants. Baumol (1952) refutes the classical economist's notion that the transactions demand for money depends solely on the level of income. This conclusion evolves from the following assumptions. Receipt inflows, which occur periodically, arrive in the form of bonds while payment outflows follow a continuous steady stream over some time period. Moreover, firms know with certainty the current and future pattern of these flows. Finally, two assets--bonds and money--can serve as stores of wealth.

Given these conditions, the value of a cash balance arises from the need to smooth out the unsynchronized receipt-payment schedules. While only money can perform this task, firms prefer to keep wealth in interest earning form until payments come due. Yet, holding bonds requires making a transfer payment when converting into spendable cash. Thus, using money for transaction purposes entails a two-part cost; conversion costs of b(T/C) and an interest payment foregone to hold cash equal to i(C/2). The symbols have the following definitions: T = payment stream, C = amount of one transfer from bonds to cash, i = opportunity cost of holding money, and b = opportunity cost to convert bonds to cash.

A rational business desires to minimize this cost by choosing the proper size of its control variable (C). The resulting optimum cash withdrawal implies optimum average money holdings as,

$$M = \frac{C}{2} = \left(\frac{bT}{2i}\right)^{1/2}$$

which states that the transactions demand for money varies directly with transfer costs and the flow of transactions and inversely with changes in the interest rate. Finally, the function's most celebrated implication is the existence of economies of scale in money holding with respect to transactions flows.[1]

Reaction to Baumol's findings appeared rapidly. Theoretically, the discussions focused on the model's rigidity with respect to both its assumptions and conclusions. Some critiques resulted in support of the simple framework, while others rendered contradictions.

Tobin (1956) reaffirms the conclusions made in the Baumol study as a special case of a more general setting. Starting with the Baumol framework, he allows for both fixed and variable transfer costs. In addition, firms choose bond and cash holdings to maximize interest

---

[1]These two properties contrast the Keynesian postulate that the transactions demand for money moves proportionally with income.

earnings net of transfer costs. Given this goal finding the optimal amount of money to hold requires three steps.

First, he derives the optimum transfer schedule, which has two important characteristics. All conversions from cash to bonds must occur at the start of each period to avoid losing interest payments which justifies Baumol's assumption that receipts come in the form of bonds. Also, any transfer from bonds to cash must occur when cash balances fall to zero; otherwise, total interest earnings will decline. This finding supports the simple sawtooth process implied by the Baumol model.

The second step determines the number of transfers necessary to maximize profits. Four cases exist depending on the magnitude of interest rates and transactions levels relative to transfer costs.[2] The Baumol model applies only when transactions and rates are large relative to transfer costs. Such circumstances apply to firms, inducing them to hold balances of both bonds and money and make numerous switches between them.

Given this Baumol case, the third task proves an inverse relationship between the interest rate and money holdings.[3] In addition, cash balances vary directly with transactions amounts and transfer costs. Thus, the Tobin analysis encompasses the Baumol model as a special case especially applicable to firms.

While this work supports the Baumol conclusions, other critiques do

---

[2]This result stems from Tobin's (1956) application of inequality constraints on the number of transfers.

[3]This conclusion comes from maximizing interest earnings net of transfer costs, provided the result exceeds zero. If not, firms will hold no bonds.

not. Brunner and Meltzer (1967) prove that, under certain conditions, no inconsistency exists between the Baumol model and the quantity theory. The conditions include; the presence of different marginal transfer costs into as opposed to out of money and positive cash flows from sales as well as bond transfers.[4] The latter property allows firms to acquire money without taking it from bond holdings. Thus, total money holdings equal the weighted sum of optimal amounts from both sources and depend on transactions levels, the interest rate, and all transfer costs. An explicit form of this function has the Baumol model as a special case where marginal switching costs equal zero. With positive variable transfer costs and no fixed costs, the equation reduces to the quantity theory. When all transfer costs exceed zero, the equation's elasticity coefficients do not take fixed values. This property implies that estimation of an inventory model should not constrain these values.

Another route to the same conclusions comes from Sprenkle (1966) by incorporating demand deposits into the Baumol model. Since banks pay positive returns on such balances, firms try to maximize total earnings from bonds and deposits net of transfer costs. The resulting money demand function implies larger holdings than the Baumol form. In addition, he argues that payments on demand deposits vary directly with deposit size (and, thus, with transactions amounts) and bond interest rates. These relationships alter the Baumol model elasticity values.

Specifically, the transactions elasticity will exceed one-half since changing deposit rates move money holdings in the same direction as

---

[4]Baumol (1952) also considered this case but did not formally analyze it.

changing transactions. Similar reasoning implies that the bond interest rate elasticity has an absolute value less than one-half. While leaving the original variables intact, the Sprenkle analysis makes their elasticity values econometric questions.

Barro and Santomero (1972) extend the notion of splitting money into currency and demand deposits. They dichotomize money holdings into the two forms by explicitly including a positive return on demand deposits. In addition to this change, a fixed transfer cost exists between currency and demand deposits which is less than the fixed cost of changing bonds into money.

Given these characteristics, individuals try to maximize interest earnings net of transfer costs from managing the three assets. The solution has optimal money balances changing directly with transactions and the difference between transfer costs from deposits to bonds and those from cash to deposits. Finally, money balances are inversely related to the difference between the rate on bonds and that on deposits.

Barro (1976) took the model developed by Barro and Santomero and altered it to include integer constraints on the number of transfers between assets.[5] This addition results in a step function where individuals, attempting to maximize net earnings, alter their behavior when reaching certain thresholds. This result causes an aggregation problem since the sum of individual action will give a continuous function in all variables.[6]

---

[5]This process mirrors Tobin's (1956) addition to Baumol's (1952) original work.

[6]According to Barro (1976), this conclusion rests on a diverse distribution of individuals.

In particular, individuals' income can be accurately described by a first order gamma distribution which implies that aggregate expenditure decline geometrically during a single period. Given this property, the optimal number of transfers (an integer) depends on the expected value of expenditures, transactions levels, the difference between the rate on bonds and money, and transfer costs. Money demand depends inversely on the optimal number of transfers. In addition, elasticity values can take a range of values. The elasticity of transactions falls between one-half and one while that for the opportunity cost of money has a range from a negative one-half to zero. Finally, the elasticity with respect to transfer costs lies from zero to one-half. As the average number of transfers goes to infinity, the elasticities take the values predicted by the Baumol model. Using simulations, Barro shows that if the average number of transfers exceeds one and one-half then the elasticity values from the Baumol model will lie very close to those predicted by the integer constrained model. This conclusion supports unconstrained estimation of deterministic inventory equations.

The previous two articles add a positive rate on demand deposits to an integer constrained inventory model to derive their results. Another addition to the Baumol-Tobin model makes commodity purchases an endogenous variable.

Feige and Parkin (1971) broaden the Baumol-Tobin analysis to allow endogenous expenditure levels by incorporating the fact that cash management decisions affect real receipts available for expenditures.

Specifically, individuals try to maximize utility from expenditures subject to a budget constraint that includes the net revenue from managing money, bonds, and goods. In contrast to the Baumol model, individuals must now consider the opportunity cost and returns of holding all three assets as well as the transfer costs between them. The solution requires discovering the optimal number of transfer between money and bonds and also between money and goods. The resulting optimal cash balances depend directly on the rate on money, total expenditures, the transfer cost between money and bonds, and the holding cost of bonds. The rate on bonds and the holding costs of both money and goods affect cash balances inversely.

This basic analysis was merged with that of Barro and Santomero (1972) by Santomero (1974). He adds two characteristics to the Baumol model. First, money consists of two distinct types of debt: currency and demand deposits. Total money demand must equal the sum of the holdings of each type. Second, a non-zero cost exists to transfer money into commodities. This addition implies a finite number of commodity purchases which rationalizes holding an inventory of goods and influences money demand.

Within this framework, the individual has four ways to hold receipts: currency, demand deposits, bonds, and commodities. He attempts to maximize the net return on the four balances which depends on the return to each asset and the transfer cost between them. This task is accomplished by controlling the number of transfers between bonds and money, currency and demand deposits, demand deposits and goods, and currency and goods. The resulting optimum amount of money holdings depends directly on the volume of transactions, the difference between the transfer cost of money to bonds compared to that for demand deposits to currency. An inverse relationship exists for money holdings and the difference between the rate on bonds and that on deposits and the optimum inventory of goods. This function contains the Baumol model as a special case where zero cost exists when buying goods. Non-zero costs will result in smaller money holdings since individuals keep positive commodity inventories. Finally, money demand depends directly on the differences between the rates on money (either form) and goods and negatively on the difference between rates on bonds and deposits.

Grossman and Policano (1975) take the Santomero model one step further by allowing the purchase of a good less frequently than the receipt of income (a capital good). This addition explains positive bond holdings as a way to accumulate funds to buy capital not as an alternative to money. Bonds will not serve in the latter capacity when transfer costs are large relative to interest earnings.

They postulate the existence of money, bonds, goods, and capital each having a positive return and a fixed transfer cost. The individual determines the frequency of transfers (and, thus, the average holdings of each asset) to maximize interest earnings net of transfer costs. The resulting patterns of transfers imply a two-part money balance. The working balance is used to buy goods while the saving balance accumulates for capital purchases. Over one period the former declines uniformly to zero as the latter increases uniformly from zero. Furthermore, the working balance depends inversely on the transfer cost of and return on goods and the level of goods inventory. The saving balance is directly related to the inventory of capital and the transfer costs and return to capital. It also varies inversely with the rate on savings.

Each of the studies outlined after Tobin (1956) suggests additional variables to include in money demand. Yet, lack of data precludes implementing the suggestions in empirical work. Their major empirical implication postulates estimation unconstrained by the square root law.

Yet, they do not represent a radical departure from the deterministic inventory approach and, therefore, can easily be incorporated into the B-T formulation. A different line of criticism developed which substantially alters the conclusions coming from an inventory model. The main feature of these critiques is the replacement of deterministic receipt-payment flows with stochastic patterns.

Stochastic Variants. Tsaing (1969) faults the Baumol-Tobin model for assuming certainty in expenditure patterns. If a non-zero probability of unforeseen expenditures exists, holding precautionary cash balances becomes rational. The liquidity of these holdings prevents conversion costs (from bonds) or postponement of purchases when unforeseen expenses arise. Thus, the yield on precautionary balances equals the expected loss due to forced emergency sale of assets to meet unplanned purchases.

Adding these balances to the Baumol model, individuals try to maximize the return on managing assets net of transfer costs. The solution requires finding the optimal precautionary and transactions holdings simultaneously. Both types depend on transactions, transfer costs, the rate on bonds, the expected loss function, and the probability density function of expenditures. Finally, imposing integer constraints allows solving for the optimal number of transfers. This number (and, thus, money holdings) are insensitive to changes in the variables listed above. For example, transactions must increase threefold before the number of transfers changes from two to three.

While providing a rationale for incorporating uncertainty into money demand, the mainstream of stochastic models stem from analysis by Miller and Orr. Miller and Orr (1966) derive the implications of an inventory model for firms' money demand when receipt-payment schedules follow a random pattern. They incorporate the uncertainty faced by business under the following assumptions. Two assets exist: cash and treasury bills, which yield (r) per day. Switching wealth from one form to the other costs b per transfer. Net cash flows are generated as a sequence of independent Bernoulli trial where each change in cash (m) has an equal probability of being positive (p) or negative (q). An asymptotically normal probability distribution results with a mean of nmt (p - q) = 0 and a variance equal to $4ntpqm^2 = nm^2t$ where n represents the number of days and t is the number of trials per day. Thus, the average daily change in cash (σ^2) equals m^2t and measures the level of uncertainly about cash flows.

Given these assumptions, firms attempt to minimize the expected daily cost of managing money balances as the sum of the expense for making the expected number of transfers plus the interest earnings foregone on mean daily cash balances. The optimal control rules allow cash levels to change randomly until reaching either an exogenously determined lower bound or an upper bound. If money balances break out of this range (h), firms should restore them to a level z above the lower bound. Imposing these rules to control expected daily costs results in a mean cash balance of (h+z)/3. Solving for the values of h and z which minimize expected cost and substituting into the formula for mean cash balances gives the firm's money demand as,[7]

---

[7]The elasticity values of one-third stem from the fact that the optimum value for z lies one-third of the way from the lower to the upper bound.

$$\overline{M} = \frac{4}{3}\frac{(3m^2t)}{4r}^{1/3} = \frac{4}{3}\left(\frac{3b}{4r}\sigma^2\right)^{1/3}$$

which depends on transfer costs (b), the interest rate on bonds (r), and the variance of daily cash flows (σ^2). Even though the analysis represents a significant departure from deterministic inventory models in both assumptions and conclusions, it rests upon some unrealistic postulates. The sensitivity of M-O conclusions to changes in assumptions is critical to this model since empirical work must wait until data become available for the variance of daily cash flows.

Miller and Orr (1968) show the insensitivity of their stochastic model to changes in the Bernoulli cash flow process, the assumption of only two assets, and the existence of only fixed transfer costs.[8] First, their equation yields accurate results even when the cash flow distribution has a stable positive cash flow drift (an expected value exceeding zero) or if it follows a normal distribution with a standard deviation small relative to the boundary range (h) for cash balances.

Second, the existence of two alternative assets to money does not alter their money demand equation since shorts (short maturity asset with low transfer costs) serve as sole recourse for transfers into and out of money. Balances of longs (long maturity assets with high transfers costs) are dichotomized from the decision about cash holdings. This property allows firms to control money as if only two assets exist and, thus, money demand remains identical to the original form. Finally, Miller and Orr show the original model invariant to three different types of transfer costs. Replacing fixed with proportional transfer costs changes the

---

[8]The first two criticisms come from Eppen and Famma (1969); the third assumption was critiqued by Weitzman (1968).

elasticity coefficient prediction to the original model. Finally, different constant costs for transfers into as opposed to out of cash do not alter any of the elasticity values of the original model. These additions imply that estimation of the stochastic inventory model should occur with unrestricted coefficients.

This same conclusion stems from incorporating demand deposits into the Miller-Orr model. Frost (1970) forms a money demand function by including the market for banking services. He notes that the Miller-Orr model describes discretionary money balances above some exogenously determined lower bound. Actually, this minimum amount--compensating balance--depends on interaction between banks and firms for banking services. Thus, total discretionary balances equal the sum of transactions holdings, as predicted by the Miller-Orr equation, and compensating balances.

Given these circumstances, firms attempt to maximize the earnings net of costs of total discretionary balances, subject to the banks' supply of services. When that supply depends on the level of transactions balances (and, therefore, on the firm's control variables) the levels of compensating and transactions holdings must be determined together. If not, firms will adjust deposits too often, hold excessive compensating balances, and receive too few services. Finally, the interaction between the two holdings makes the money demand elasticity magnitudes econometric questions.

Summary. Basing money demand on inventory theory arose as a microeconomic response to the Keynesian assumption that transactions demand for money depends only on income. In contrast, inventory theory describes the factors influencing any rational agent's cash balances when alternative stores of wealth exist. Explicitly incorporating these assets implies that demand for money will depend not only on transactions levels but on the benefits and costs of using all assets. Furthermore, these additions destroy the proportionality predicted by Keynesian theory when implying economies of scale in transactions money holdings.

An exact relationship from inventory theory has been developed along two different lines.

The pioneering work of Baumol and Tobin and subsequent extensions based optimum cash holdings on deterministic operating cash flows. Using this knowledge, firms alter cash balances in response to changes in transactions levels, the cost of transferring wealth between assets and money, and the interest earnings from holding assets. The B-T model predicts elasticity values for these three determinants as a positive one-half, one-half, and a negative one-half, respectively. By altering the assumptions of B-T, other authors derive money demand where these elasticities values are econometric questions. However, they do not alter the basic form of money demand.

In contrast to deterministic variants, Miller and Orr originated a version on a random pattern of operating cash flows. Firm's money holdings depend on the variance of daily cash flows, transfer costs, and interest earnings. The M-O approach predicts these variables' elasticity values as one-third, one-third, and a negative one-third, respectively. Using broader assumptions, subsequent work gives no specific magnitudes to these elasticities a priori. Finally, the simple M-O model appears robust with respect to changing simplifying assumptions.

Empirical Work

Overview. Inventory models stand alone in predicting scale economies in cash holdings. This feature coupled with the conclusion that money balances are sensitive to changes in interest rates and transfer costs make for a radical departure from the classical notion of money demand. Consequently, empirical work centered on attempting to prove or disprove their existence. The earliest tests examined the absolute ability of inventory models using indirect methods while later studies employed direct estimation. Finally, attention turned to using direct estimation in order to compare alternative models. Since the literature has advanced beyond the indirect method, only absolute and comparative tests appear below. For examples of tests using indirect means, see McCall (1960) and Seldon (1961).

Absolute Tests. Although employing different techniques, each direct test attempts to verify or reject one theoretical form with some body of data. Two studies testing Baumol's equation appear first followed by an estimate of the M-O variant.

Ben-Zion (1974) finds support for the cost of capital (r) as the proper interest rate in a inventory equation. He constructs this rate as $r = \frac{E}{P} + h$, where P is the price of a corporate share, E stands for earnings per share, and h represents the long run growth of E. Substitution of r into the Baumol model gives a specification which he tests using one observation from 546 firms for each variable in the year 1964. The results support the specification and imply economies of scale. While the use of direct estimation is an important aspect of the Ben-Zion study, it also highlights the liberal interpretations of money's opportunity cost, an important topic in money demand tests. Finally, even though the results support an inventory equation as an explanation of cash balances, the model may not make accurate predictions.

Sprenkle (1969) shows that an inventory model makes inaccurate predictions of actual money balances held by firms. For the year 1968, he assumes a five percent interest rate and transfer costs equal to 20 dollars per switch. Substituting these values along with sales levels for 475 firms into the Baumol equation results in predictions of the firms' money balances. Since these predictions represented only two and one-half percent of actual cash holdings, Sprenkle rejects the Baumol equation. He argues that the unexplained portion consists of compensating balances, not transactions balances making inventory equations useless as descriptions of behavior. In addition to highlighting the problem posed by compensating balances, the Sprenkle analysis indicates the importance of using prediction to judge models.

Orr (1971) employs the same technique to prove that a stochastic inventory model can predict accurately. Using Sprenkle's values for the interest rate and transfer costs, he forms predictions of money demand from the Miller-Orr equation. Figures on the variance of daily cash flows were constructed from cash flow data over a 300-day period for three firms. The best prediction accounted for 80 percent of actual holdings, while the worst prediction was twice as accurate as Sprenkle's best estimate. Although proving his point, the study shows the data availability problem when testing a stochastic inventory model.

While absolute tests of single equations provide much useful information about a model's feasibility, an unusual sample could cause the rejection of a correct model. Comparison tests can greatly supplement single equation regressions by providing relative measures of the strength of particular models.

Comparison Tests. Studies in this category use a common body of data to estimate alternative specifications designed to explain money demand. Only one such paper appears here; others are discussed after the development of the relevant alternative theories.

Whalen (1965) compares the Cambridge quantity theory and the Baumol inventory model as descriptions of firms' money demand. He derives testable forms of the two as,

$$\frac{M}{MA} = a_1 + b_1\left(\frac{S}{MA}\right)$$

$$\frac{M}{MA} = a_1 + b_2\left(\frac{S}{MA}\right) + C\left(\frac{S}{MA}\right)^{1/2}$$

respectively, where M stands for money balances, MA represents monetary assets (cash plus short term sercurities), and S symbolizes sales. He employs data from eight industries, each stratified into 14 size classes. Using all observations, estimation supported inventory theory and implied economies of scale. Yet, the results from individual industry regressions rejected the inventory equation in seven cases. Since the slope coefficient of the Cambridge equation differed among industries, Whalen rejects it as a general description of firms' behavior. This study points out the importance of comparing models as alternative explanations. It also highlights how results can change depending on the level of aggregation used.

Summary and Criticisms. The empirical literature does not provide conclusive evidence about the validity of inventory models. This failure stems from three shortcomings contained in the group of studies. Correcting these faults requires application of the three part econometric methodology to money demand theories. Each part corresponds to a common shortcoming of the empirical work.

First, since inventory theories explain individual firm behavior, they should be tested using firm level data. The studies of Ben-Zion (1974) and Whalen (1965) employ aggregate data without addressing the potential problems. Aside from being logically inferior to using firm level data, this method only gives accurate results in the absence of aggregation bias. Measuring the amount of such bias and testing its significance requires individual firm data.

Second, while inventory theory generates testable specifications, the empirical literature has not consistently employed them. Although transactions, the interest rate, and transfer costs jointly determine cash holdings, only transactions is contained in every study. Even here no consensus exists as to the relevant proxy for transactions. Ben-Zion (1974) and Sprenkle (1969) employ nominal sales while Whalen (1965) uses sales as a percent of assets. Yet, accounting data gives entries that mirror changes in transactions more accurately than these variables. With respect to an interest rate and transfer costs, Whalen (1965) omits them both from his equation. While Ben-Zion (1974) includes an interest rate variable, his equation excludes transfer costs. These misspecified models give unreliable results about inventory models. Finally, only Sprenkle (1969) and Orr (1971) incorporate the crucial role of transfer costs in inventory models.

Third, these studies do not use the most powerful econometric techniques available. For example, while Sprenkle (1969) and Orr (1971) highlight prediction as an evaluation tool, they do not apply the most robust method in constructing predictions. Their technique fixes the equation's parameters and then substitutes actual values for the independent variable to generate values for the dependent variable. Calculating such "fitted" values and comparing them to actual dependent magnitudes is the function of least squares estimation. Such a method makes minimum error estimates about actual values by calculating parameter magnitudes during the "fitting" process. This analysis not only shows that least squares give superior results, it also implies that the two studies do not actually make forecasts. The objective of predictions is to accurately describe future outcomes. While the remaining studies use least squares estimation, circumstances can arise requiring more powerful variations. Both undesirable disturbance properties and pooling data represent such special cases.

One function of Chapter III is to rectify these three deficiencies. Gathering a proper sample, employing advanced econometric techniques, and forming specifications are each discussed in detail to ensure reliable comparative tests.

Portfolio Model

Theoretical Work

Overview. Money demand can be modeled in a manner similar to the demand for any durable good. Money belongs in this category because holders derive utility from the flow of services it provides. Yet, several alternative stores of wealth exist, each having some unique, desirable property. Faced with this selection, a rational wealth holder arranges a portfolio containing a variety of such assets. For any given amount of wealth, the optimum level of each asset is determined by equating the marginal rates of return on all forms. The resulting demand for money depends on total wealth, the price of money, and the price of substitute assets.

Friedman (1956) uses the quantity theory to derive money demand as a function of wealth, the return on money, the rates of return on assets that substitute for cash, and the investor's preferences.

Wealth (W) represents the stock accumulating from the discounted flow of retained earnings. To incorporate the essence of the portfolio approach, Friedman considers four stores of wealth: money, bonds, equities, and physical goods. Money represents a medium of exchange with a fixed nominal value. Its return stems from these two asset characteristics as well as security, convenience, and pride of possession. Bonds are claims to a stream of payments of a fixed nominal amount per period. These earnings plus any capital gain or loss due to change in the bond's price make up the return received by holding bonds (r_b). Equities render a share of profits from an enterprise and pay a three-part return (r_e): a fixed payment per period and a capital gain or loss due to a change in either the fixed payment or the general level of prices (P). Physical goods yield a capital gain or loss when the general price level changes. Finally, preferences (U) determine the form of money demand.

Combining these considerations gives,

$$M = F\left(\overset{+}{P}, \overset{-}{r_b}, \overset{-}{r_e}, \overset{-}{\frac{1}{P}\frac{dp}{dt}}, \overset{+}{W}, U\right)$$

as money demand. Given this function, the optimum cash holdings occur at the point where wealth is allocated to the four assets such that their marginal returns are equal. This rule maximizes the total return from the portfolio and establishes the partial derivative signs appearing above each variable.

Although Friedman forms the analysis for a household, he contends that it applies to firms in the same form. The amount of money a firm holds depends on the cost and value of productive services from money as well as the cost of substitute productive services. These will be outlined in turn. The cost depends on how the corresponding money holding is raised: by bonds or equities, by substituting cash for real capital goods, etc. These ways of financing money balances correspond to the alternative form in which households can hold wealth. Thus, the rates of return in the money demand equation represent cost to business of holding money.

The value of money's productive services depends on the production function. It will be sensitive to factors affecting the smoothness and regularity of operations, size, degree of vertical integration, etc. Since none of these variables merits special status in the money demand function, Friedman incorporates them into U along with tastes.

Finally, substitutes for money as a productive service include many ways of economizing on money balances by using other resources. Specifically, they can help synchronize payments and receipts, reduce payment periods, extend book credit, etc. Since no particular close substitutes exist, their prices do not need to appear in firms' money demand.

The combination of these considerations leaves business money demand with the same form as that for households with an expanded interpretation of U.

Summary. The general nature of Friedman's equation allows it to encompass any other work along the same lines. Since its conclusions do not depend on questionable assumptions, no major extensions of the general framework exist on a theoretical level. The resulting money demand equation embodies two unique aspects of portfolio theory. First wealth acts as a constraint on holding additional money. In contrast, inventory models employ transactions levels to measure the need for money. Furthermore, they postulate a fixed factor relationship where making a dollar's worth of payments requires a certain amount of money. The portfolio approach treats this relationship as a result of the utility maximizing process (Friedman, 1956). For example, if money's holding cost increases, rationality implies reducing the volume of transactions per dollar of cash balance. The optimum relationship between money balances and transactions depends on the rate of return on an alternative asset since it measures money's holding cost. Second, the portfolio approach includes a broad scope of alternatives to holding money. Several rates of return must appear in the equation to capture these substitution possiblities.

Empirical Work

Overview. Although Friedman's equation rests on sound theoretical ground, its generality poses two empirical problems. Multicolinearity may arise with the inclusion of many rates of return. Also, no specific functional form is implied by the analysis. Empirical studies must deal with both of these problems before estimation by forming a testable specification. After justifying their own form, the studies use direct estimation of either a single equation or comparisons of a set of alternative money demand forms. These two approaches will be explored in turn.

Absolute Tests. Meltzer (1963a) finds support for a money demand specification formed from the Friedman equation by using the following assumptions. The function has first degree homogeneity in wealth (W). Furthermore, sales (S) and wealth have a proportional relationship as S = KW. The value of K depends on the internal rate of return, capital-labor ratios, and asset utilization rates. A single rate of return (r) measures the yield on bonds, equities, and physical assets. Since K and r remain constant for a cross section of firms, Meltzer combines them into a single term (V). Finally, money demand has a log linear form as

$$M = VS^b$$

which he tests with a 14 industry sample. For each industry, 10 asset sizes exist which have data on M and S.

Estimation for each industry rejects the existence of either economies or diseconomies of scale. The sale elasticity of one supports the quantity theory as an approximation of money demand. This study gives an example of how the literature dealt with the general nature of the portfolio equation.

Testing of single equations provides only absolute measures of a model's ability. Comparison tests, employing alternative equations as descriptions of money demand, expose a model's relative capability.

Comparison Tests. Each study in this group allows different specifications to explain firms' money demand.

Whalen (1965) compares the Meltzer equation with a specification formed from that equation as follows. He argues that r and K vary directly with firm size and, thus, do not stay constant even for a cross-sectional study. Adding these relationships to the Meltzer analysis gives money demand as,

$$M = a\left(\frac{W}{S}\right)^{b} S^{c}$$

Whalen compares this model to the Meltzer equation based on a sample containing one observation for the 14 size classes of eight industries in the year 1958.

Estimation supports the Whalen equation since the ratio of wealth to sales had a significant coefficient. This result implies that the Meltzer equation gives biased sales elasticity estimates. Using the Whalen specification supports the existence of economies of scale. This conclusion highlights the importance of forming a correct specification to represent a theory.

In another comparative study, Vogel and Maddala (1967) prove the importance of using advanced econometric techniques in estimating money demand functions. They employ two samples in distinguishing between inventory and portfolio theories. One contains cross-sectional data for 16 industries tiered into 14 asset size classes in the year 1960. The other has time series observations from 1947-1961 for total manufacturing stratified into 10 asset size classes.

The equation used to test for economies of scale appears as, LN (cash) = a + b LN (sales) + DUMMIES. For the cross-sectional data, the dummy variables adjust for industries and asset size classes. When using the time series sample, dummies are inserted for each year and every asset size class.

When pooling the industry data, Vogel and Maddala found the following. Both industry and size class dummy variables had significant coefficients. Inclusion of the size class dummies significantly lowered the estimate of the sales elasticity. Thus, omitting the influence of assets assigns too much explanatory power to sales resulting in the rejection of economies of scale. Identical results occurred with the time-series sample. These findings highlight the importance of using advanced econometric techniques in money demand studies.

Summary and Criticism. These studies do not give conclusive results about the portfolio approach. Three deficiencies exist that contribute to such a result. First, portfolio theory describes individual firm behavior; yet, the studies use aggregated data without addressing the aggregation problem. Dealing with the potential aggregation bias requires firm level data.

Second, forming a portfolio specification requires addressing two factors. For one, the specification for money demand contains two unique aspects. Capturing its essence requires the use of wealth as a scale factor and the inclusion of many rates of return to measure substitution possibilities. Without these factors, a portfolio specification will not give testable contrasts to alternative models. Meltzer (1963a) uses sales as a scale factor and only one rate of return. Vogel and Maddala (1967) also employ sales but omit all returns on alternative assets. While Whalen (1965) omits all rates, he uses sales as a percent of assets for the scale variable. The studies' avoidance of using wealth stems from trying to provide comparable results to previous studies that employed sales to generate elasticities and test economies of scale. Yet, the portfolio equation employs wealth purposely to measure the constraint on holding money. Since widespread data exists, wealth should appear in the portfolio equation. The studies surveyed justify the omission of many rates of return on the basis of potential multicollinearity. Yet, ridge regression can deal with this problem. Also, the portfolio equation takes a general form, forcing the selection of a testable specification. All the studies surveyed assume a log linear form without comment on possible misspecification. The Box-Cox technique can provide optimal selections.

Finally, accurate estimation may require more powerful methods than ordinary least squares. The existence of autocorrelation, heteroscedasticity, or other undesirable circumstances necessitates the use of such methods.

Chapter III contains detailed analysis of each of these faults. In combination, they form a methodology which renders accurate model evaluations.

Production Model

Theoretical Work

Overview. Money demand can be modeled as an extension of the productivity of money. A variety of reasons exist to explain how cash balances lower firms' real costs of production. In each case, the implication remains the same: money holdings must explicitly enter as a productive element in the profit maximization process. The resulting demand for money is derived from output demand and depends on the value of output, the price of money, and the prices of physical inputs.

---

[9] The basis for their analysis comes from Lange (1936).

Transactions Cost Variant. Gabor and Pierce (1958) outline the transactions cost view of money demand.[9] They note that prior to any production run, firms require time to secure and ready inputs. Despite this lumpy build-up process, businesses wish to establish an even flow of ouput. Money provides a means to smooth the process by giving the firm command over physical resources without having to hold them in inventory prior to use. Since inputs are hired for shorter periods of time, real payments to their owners decline for any production run.

In spite of money's productivity in lowering real costs, it remains fundamentally different from physical resources. Production could occur without money; yet, it requires every physical input. Thus, cash holdings influence the use of physical resources from outside the production function. In other words, the firm's decision process is dichotomized into a physical side and a financial side, inducing firms to maximize profit subject to both a production function and a money requirements function. The latter relates the dependence of cash balances on prices and levels of output and physical inputs.

At any point in time, money requirements equal the difference between the discounted values of receipts from output and payments to inputs implying that necessary cash depends on the price and quantity of inputs and output and the interest rate. Gabor and Pierce do not derive money demand since their interest rested only with the money requirements function.

Saving (1972) generates an explicit form for firm's money demand within the transactions cost framework. He begins with a production function showing the dependence of output flows on the levels of physical inputs. Normally such a relationship either implicitly includes or ignores the transactions costs involved in acquiring inputs and selling output. While this simplification creates no problems for most analyses, these costs are of primary importance for the study of firm's money demand. Thus, they must appear explicitly in the decision process. Such a function shows money requirements dependent on the dollar value of both output and physical inputs and the average holdings of output, physical inputs, and money. Given an explicit form of both the money requirements and production functions, firms maximize the present value of the firm's profit stream. The resulting solution for money demand becomes,

$$M^* = g(\overset{+}{P}, \overset{-}{r}, \overset{+}{H_X}, \overset{-}{H_V}, \overset{-}{H_M})$$

where r is the discount rate and H_X, H_V, and H_m stand for the holding cost of output, inputs, and money stocks, respectively. Partial derivate signs appear above each variable and state that M* varies directly with the price of output and the cost of holding an inventory of output and indirectly with interest rates and the costs of holding both physical inputs and money. Similar results stem from an alternative view of both the reason for and the proper way to model money's productivity.

<u>Neoclassical Variant</u>. Johnson (1968) provides an explanation of money's productivity in a macroeconomic context. He begins by assuming that production in a barter economy requires labor and capital. Furthermore, capital takes two forms; productivity equipment and inventory. Given these conditions, the economy can produce a maximum level of output by correctly combining the three inputs. Assigning one capital good (X) as a medium of exchange alters this optimal input combination.

Firms can now substitute the holding of a smaller value of X for other capital goods in inventory. Such action results in production of the same output level using smaller amounts of both types of capital. Furthermore, the introduction of a fiat money allows firms to use their entire stock of inventory capital in production by substituting cash for it. The resulting shift of the production function reflects money's marginal product. Since the same process occurs for changes in physical inputs, cash enters the firm's production decisions as any other resource. This analysis provides only the skeleton of the neoclassical view because its formation occurred on a macro level. The missing substance is contained in a microfoundation which accounts for individual firm motivation.

Moroney (1972) outlines the reasons that firms hold inventories of goods and substitute money for them. The introduction of a fiat money into a barter economy benefits firms in two ways. First, money lowers the real cost of production since resource owners will accept less money vis-a-vis final goods in payment for a given exchange of inputs. By eliminating the double coincidence of wants problem, money increases the holders leisure time. Thus, the total compensation paid--money plus extra leisure time--remains equal to the in-kind payment, while the firm's share--money--declines. Second, with barter, supervision and delivery of in-kind payments use human time and capital goods. Money, by facilitating the use of credit and contracts, lowers the amount of real inputs needed to carry out these tasks. For both reasons, the

---

[10]Moroney (1972) does not believe that the firm can control these combinations. Hence, he models money as an external technological improvement, not as a production function input.

innovation of money increases the number of alternative combinations of resources available to product output.[10] Later studies began to derive mathematical versions of this view.

Levhari and Patinkin (1968) place cash explicitly into the production function and describe how marginal productivity theory determines money demand. They postulate a production function showing aggregate output dependent on the amounts of labor, capital, and money. Maximizing profit subject to this function requires equating the marginal advantage of all inputs. The joint solution of this expression and the total expenditure constraint gives the derived demand for money. It shows cash balances dependent on the value of output, the interest rate, and the prices of capital and labor.[11]

Both the transactions cost and neoclassical variants can provide production specifications. Fisher (1974) outlines the relationships between the transactions cost and neoclassical variants under the following assumptions. Firms have a physical input production function where no transactions costs exist to buy inputs or sell output. Net receipts from operations can be held as cash or bonds. Finally, transfers between the two entail a fixed charge.

Under these conditions, the firm maximizes profits in two steps. First, the input-output combination that creates the most net receipts is found. The buying of inputs and selling of output depends only on the physical aspects of the firm and generates a non-coincident receipt-schedule. This pattern motivates firms to hold money balances in addition to bonds. The second step determines the optimal amount of net

---

[11]They do not provide a microfoundation of money demand since their concern lies with economic growth.

receipts to hold in each of the two assets. Such a decision deals only with the financial aspects of the firm.

Although total profit depends on money holdings, the absence of interaction between the two steps in the decision process precludes modeling money as an input. To create interdependence, Fisher links transfer costs to labor. He assumes that one transfer uses the services of one unit of labor, causing output to depend directly on money balances. Any increase in cash holdings results in additional transfers leaving less labor for production.

Fisher establishes a synthesis between the two production variants; however, the analysis does not continue on to derive money demand functions. Finnery (1980) outlines a mathematical approach to this topic under the following assumptions. A physical production function exists which shows different output flows (Q) produced with technically efficient combinations of capital (K) and labor (L)--$Q = g(K, L)$. For each output level and input pair, a pattern of cash receipts and payments results which together determines the minimum level of money balances (M) needed to facilitate all transaction--$M = m(Q, K, L)$. Using these transactions cost conditions, the firm attempts to find the minimum cost combination of K, L, and M subject to the physical production function.

As an alternative, the neoclassical position occurs as follows. With regard to the money requirements function, if a change in Q alters M, the implicit function theorem implies that $Q = h(K, L, M)$. This result means that output can be described by a money inclusive production function as $f(K, L, M) = g(K, 1) + h(K, L, M)$ which will give equivalent answers to the cost minimum problem as using the g and m functions combined. Thus, the two methods give identical money demand equations. This synthesis lends further support to employing a money inclusive production function to model firm's behavior when they require money for transactions.

Summary. Money as a productive input can be modeled in two ways. The transactions cost variant stresses money's difference from physical inputs. The resulting decision process contains two parts: a physical side represented by a production function and a financial side modeled by a money requirements function. In contrast, the neoclassical variant emphasizes the similarities between money and physical inputs. Firms make decisons subject to a money balances inclusive production function. In either case, the resulting money demand equation stems from the demand for output and depends on the prices of all inputs as well as the value of output.

Such a formulation contains two unique characteristics. First, physical output represents the scale variable. It measures the cash necessary, in combination with other inputs, to produce at minimum cost. Second, the inclusion of resource prices incorporate the opportunity available to produce with different input combinations.

The production approach sees money demand as a result of using it with other resources to obtain the greatest total output possible. Symmetrically, the portfolio approach views money demand coming from its use with other assets to receive the greatest total return possible.

Empirical Work

Overview. Although the conception of money's productivity differs in the two variants, the resulting money demand functions take nearly identical forms. This fact probably accounts for the failure to find any study explicitly based on the transactions cost variant.

Using the neoclassical variant, the major econometric task is to derive a specification for money demand. As Fisher (1974) points out, the resulting equation for cash balances depends on the forms chosen for other functions in the decision process as well as the process itself.

Two categories of studies exist which employ a neoclassical base. The production function studies test the plausibility of putting money into a macro production function. Money demand studies assume the viability of this inclusion then derive and evaluate money demand functions.

Production Function Tests. Sinai and Stokes (1972) test the hypothesis that money balances serve as an input in a macro production function. They assume a Cobb-Douglas form showing output (Q) as a function of technological progress (e^{at}), labor (L), capital (K), and money balances (M) as,

$$Q = Ae^{at}L^{b}K^{c}M^{d}U$$

where A is a constant and U stands for a random disturbance term.

Combining observations for total manufacturing from 1929 to 1962 with the production function and correcting for autocorrelation gave the following results. All regressors had significant magnitudes and, thus, estimation without money in the production function gives biased estimates of the coefficients for capital and labor. Elasticity estimates implied increasing returns to scale in all inputs and economies of scale in money balances. Finally, money substitutes for both capital and labor. These findings, supporting money as an input, sparked three critiques that viewed money's corrrelation with output differently.

Niccoli (1975) postulates that the S-S equation exhibits its residual pattern due to omission of aggregate demand factors not as a result of an autocorrelated disturbance. Using S-S data and adding current and lagged investment to their equation, he finds that money influences output in the S-S equation because of its correlation with investment. In a similar analysis, Ben-Zion and Ruttan (1975) argue that changes in money balances serve as a proxy for changes in aggregate demand. The latter induces innovation and technological change causing output to increase. To test this hypothesis, they add changes in money holdings to the Cobb-Douglas equation and employ S-S data. The alteration removes the autocorrelation and makes the level of money balances insignificant. The results, coupled with the significant estimate of the change in money holdings, supports their contention.

Finally, Prais (1975) also argues that the residual pattern occurs due to a specification error, not an autocorrelated disturbance. She corrects this condition by adding lagged money balances to the S-S equation. As a result, estimation with S-S data no longer exhibits the undesirable residual pattern.

Any one of these critiques precludes money as an input since they alter the production function status of the S-S equation. Yet, all the studies agree that correlation exists between money and output. The question they raise is whether cash influences output as a supply phenomena or from other behavioral relationships. As Prais (1975) suggests, the answer lies in using a simultaneous set of equations including both supply and demand influences.

Khan and Kouri (1975) add an aggregate money demand function to the S-S equation. Using S-S data, simultaneous estimation supports money as a production function input. Yet, their elasticity estimates differ in magnitude from those of Sinai and Stokes. This divergence highlights the importance of using a correct specification and leads them to conclude the superiority of the simultaneous method.

Additional support for money as an input comes from Short (1979) who compares two sets of simultaneous equations. One uses a Cobb-Douglas production function and the resulting input demand equations. The other derives a set of factor demand equations from a translog production structure. He combines these systems with data for the private business sector from 1929-1967 to render the following. The translog system reduces to the Cobb-Douglas form, justifying the use of the latter by Sinai and Stokes. The estimates support constant returns to scale in all inputs and economies of scale in money holdings. Yet, the elasticity values differed from those of S-S, indicating the importance of using the correct specification.

Substantial support exists from these papers for including money as an input. This addition implies that money demand is derived from production demand. The final two studies form such a function using different assumptions.

Money Demand Tests. Dennis and Smith (1978) prove the usefulness of basing money demand on a translog cost structure. Elasticity estimates come from applying three-stage least squares to a set of three cost share equations derived from the translog cost function. Combining the system with a sample of quarterly observations for 11 industries from 1952-1973 gave the following results.

Separability tests proved two important characteristics of the production structure. First, money balances must appear in the production structure of each industry. If not, biased estimates of the remaining inputs will occur. Second, estimation of derived demand for each input must include the prices of all other inputs, including money.

Concerning money demand, he found significant negative own price elasticities in all industries. The cross-price elasticities implied that money substitutes for both capital and labor. These findings highlight the robustness of using a translog cost structure. Yet, the results depend on the assumed production structure.

Nadiri (1969) obtains different findings from a money demand function formed with a general production structure. Firms minimize the cost of using capital, labor, and money subject to a money inclusive production function. The resulting desired money holdings depend on expected output, the user cost of capital, the wage rate, the interest rate, anticipated capital gains, and unanticipated inflation. A specific form of the function results from the following assumptions. Actual money holding adjust through a log linear process. Firms forecast expected output as a log linear weighted average of current and past output. Finally, actual money balances take a log linear form. These considerations give a money demand equation which he tests with a sample for total manufacturing from 1948-1960.

All regressors have significant coefficients, supporting the form of money demand. The interest rate figure implies an inverse relationship with cash balances. The positive estimate for physical input prices suggests that money substitutes for capital and labor. Finally, the output elasticity indicates the presence of economies of scale. While these results support the neoclassical position, the estimates contrast somewhat with those of Dennis and Smith (1978). Such divergence implies the need to use both forms in any comparative study.

Summary and Criticism. The production function tests support the neoclassical variant's inclusion of money balances in a macro production function. Thus, money demand falls under the realm of marginal productivity theory. Three topics require further discussion with respect to this process.

First, marginal productivity theory implies a microfoundation for money demand. The studies testing such functions use aggregated data without comment which is invalid unless supported by lack of bias.

Second, the production function approach allows for a variety of testable forms. Yet, no study employs the transactions cost variant as a base. Nadiri (1969) uses a general production process to derive a general form of money demand. He then assumes a log linear specification in order to conduct estimation. The Box-Cox method gives an optimal procedure to select a form. In contrast, Dennis and Smith (1976) assume a specific production process and derive a testable specification directly. Estimation of the translog cost function requires advanced econometric estimation methods. Chapter III describes gathering a proper sample, forming specifications, and using econometric methods.

Finally, these studies highlight two other topics. For one, they use real money balances as the dependent variable while the two alternative theories explain nominal cash balances. Also, Nadiri (1969) views the opportunity cost of money as having three components. Dennis and Smith (1976) use the treasury bill rate. Studies of alternative models employ a cost of capital (Ben-Zion, 1974) and long term rates (Sprenkle, 1969). To clarify these issues, a general discussion about money appears in Chapter III.

Summary

Chapter II has outlined the major theoretical and empirical studies pertaining to money demand as groundwork for the dissertation. Theoretically, the literature provides the basis of testable specifications of the three approaches to money demand. Chapter III expands on these studies to build specifications capable of making useful model comparisons. Empirically, the literature exhibits caveats whose correction ensures comparable specifications which embody the essence of the three theories. Chapter III addresses the issue of using samples containing aggregated data and circumstances requiring advanced estimation methods. The analysis leads to a methodology capable of providing useful model comparisons.

CHAPTER III

DEVELOPMENT OF THE METHODOLOGY

Introduction

Econometric analysis employs a three-part process to test theories. First, a model must be formed which describes the relationship between variables. A useful model must compromise between describing reality and being manageable. It must specify the major interrelations among variables in enough detail to provide an accurate description of reality. Yet, it must abstract from minor interrelations to remain both manageable and generally applicable. This process creates two components of an econometric equation. One part describes the deterministic influence of major explanatory variables on the dependent variable. The other component arises from omitting minor explanatory variables. It generates a random effect on the independent variable described by a probability distribution--usually normal. Thus, an econometric model outlines the probability distribution of an independent variable given the deterministic relationships.

Formally, the equation appears as,

$$y = \beta_0 + \beta_1 X_1 + \ . \ . \ . + \beta_k X_k + \varepsilon$$

where y represents the dependent variable, X_i stands for the i^{th} independent variable, and β_i symbolizes the fixed parameters expressing the relationship between y and X_i. Finally, ε is a stochastic disturbance term whose values can be considered random drawings from a probability distribution. The latter has certain properties--mean, variance, and covariance--which it projects to ε, whose presence requires applying statistical inference to estimate parameters of the model. Estimation also needs facts about the hypothesized relationships.

The second component of econometric analysis describes these facts with a set of data on the model's variables. This sample presents the relationships quantitatively so that econometric techniques can test the equation. Adequate testing requires data on all endogenous and exogenous variables, which can create a major problem. For some factors, data are not available, requiring the selection of a suitable proxy. In other cases, many alternative measures exist, forcing a difficult choice. After collecting a sample, the econometric analysis requires one final step.

This third component describes a method of transforming the data into estimators of the model's parameters. These estimators, in turn, provide a base for quantifying the model's performance. Application of the transformation to the sample gives numeric values of the estimators--estimates--which give a basis for evaluating the model.

The purpose of this chapter is to apply the econometric methodology in evaluating firms' money demand. Three distinct tasks exist, one corresponding to each part of the methodology. First, testable specifications of each theory must be formed in light of the dissertation's purpose. Not only must the equations embody the essence of their respective theory, they must allow accurate comparisons. The latter condition requires addressing theoretical questions that apply to all three theories. The issues include money's definition and opportunity cost as well as how it relates to alternative assets. Applying the conclusions from this analysis to existing equations of the three models results in comparable specifications. The existing equations represent the culmination of the theoretical work reviewed in Chapter II. Once adjusted, measuring the performance of these specifications requires combining econometric techniques with a set of facts about behavior.

Econometric methods constitute the second part of the methodology. While the least squares procedure represents the fundamental technique, each theory creates circumstances requiring the use of more powerful variants. Inventory and portfolio equations include variables that have unobservable values. Omitting them from the specifications necessitates using either the error components model or covariance model. Portfolio and production equations result in general functional forms that create two difficulties. First, a specific equation must be selected. The Box-Cox analysis performs this task in an optimal fashion. Second, the inclusion of many rates of return or prices can create multicollinearity. Ridge regression diagnoses and provides a solution to such correlation. Finally, the translog cost variant of the production model results in a system of equations. Efficient estimation necessitates a systems method such as two-stage Zellner and/or three-stage least squares. These six methods ensure statistics that accurately measure an equation's performance when mixed with a set of observations.

The third piece of the methodology deals with gathering an adequate sample. Two characteristics must exist for a data base to provide comparisons of microfoundation models. First, it must contain observations on all variables in every equation over the selected time span. Second, discovering the correct plane for estimation necessitates firm level data which allow construction and testing of aggregation bias. Results from that analysis determine the most efficient level of aggregation.

Development of the methodology occurs in the following order. First, the general questions pertaining to money are addressed. Second, for each theory, testable specifications evolve from existing functional forms given by previous work. This growth comes from adjustments in light of the discussion about money. Final testable forms result from selecting variables for the adjusted equations. Third, the advanced econometric techniques necessary to render accurate statistics of the models are outlined. The fourth section discusses the samples. Finally, a summary reviews the three-part methodology.

General Developments

Overview

Making useful comparisons of money demand theories necessitates solving general theoretical problems. Money must be defined in a way consistent with the theories. Also, all models must explain a common dependent variable. Finally, each theory shows cash balances dependent on own price. Previous literature has justified many alternatives to serve as money's oportunity cost. For consistency, one should be selected and applied to all models.

Money[1]

Money represents a medium of exchange. As such, no transfer costs

---

[1]Money is understood to mean real money balances. Putting inventory and portfolio models in real terms relies on them being homogeneous of degree one in nominal value. Since the production approach already forms money demand in real terms, adjusting the other two theories ensures that all specifications explain the same dependent variable. Thus, useful comparisons of their performance can result. Also, cash will serve as a synonym for money. Finally, the compustat tapes define money in this narrow way (see Appendix D).

exist when turning money into a spendable form. In other words, cash has perfect liquidity whereas all other assets involve some costs to be turned into a spendable form. Two types of debt perform the function of money: currency of the United States Government and demand deposits of commercial banks. This restrictive view of money leads to an analysis that justifies the federal funds rate as money's opportunity cost.

Opportunity Cost of Money

The amount demanded of any economic good depends mainly on its own price. Users of the good respond to changes in price since it represents their opportunity cost. In a market economy, the interaction of buyers and sellers establishes a set of prices which reflect the opportunity cost to consumers of alternative goods. This set provides a basis allowing rational choices among spending options. To facilitate making such choices, money provides a common denominator for evaluating relative opportunity costs. This benefit plus its function as a medium of exchange imply that money trades on one side in every market. In other words, money is a substitute for all other goods in the economy. In such a system, no direct dollar price can exist for money. To measure what the user of money foregoes, the opportunity cost principle provides an indirect way to price cash.

The principle states that the cost of taking an action equals the value of the best alternative action. Unfortunately, since money trades in all markets, every good plays the role of an alternative to money. This property explains why no consensus exists on the price of money despite the voluminous literature on the subject. The opportunity cost principle has been employed in narrowing the substitutes to only financial assets. From this point, however, an extensive list of different securities with various maturity schedules have substituted for money. Ben Zion (1974) argues for a cost of capital measure. Meltzer (1966) uses a proxy for a vector of rates. Dennis and Smith (1978) prefer a short term rate while Miller and Orr (1971) opt for a long term return. Finally, Nadiri (1969) considers money's price an econometric question and, thus, tests both short and long rates.

Properly selecting from among these alternatives requires strict application of the opportunity cost principle. Using cash means foregoing the value that would accrue from using money's best alternative. Determining this good necessitates a clear outline of money's unique attributes as an asset and then selecting the alternative that most nearly duplicates these attributes. Orr (1970) suggests a liquidity framework which can facilitate the search.

Cash has two desirable characteristics as an asset. First, it serves as a medium of exchange and, thus, has greater liquidity than any other asset. Second, holding cash entails no risk of either default or loss in nominal value. Since all assets can be described in terms of these two nonmonetary properties, they furnish a general framework for discussing money.

An asset's liquidity corresponds to the total cost necessary to transfer it into a generally spendable form. Since money acts as a medium of exchange, it resides at the top of the hierarchy which ranks all assets in terms of liquidity. The asset that nearly duplicates money's attributes resides one step away and has zero transfer costs. The farther down an asset lies, the greater the costs of turning it into money. Inducing investors to trade cash for these assets necessitates compensation in amounts at least great enough to pay the transfer costs. Thus, moving down the hierarchy involves increasing monetary returns to offset increasing transfer costs.

In fact, in a riskless world, a spectrum of returns will emerge which correspond to increasing transfer costs. While these gross returns differ due to such costs, the returns net of transfer costs will be equal at the margin. The opportunity cost of money then equals the return on a zero transfer cost asset (given a riskless world) not the higher gross return earned on an asset due to higher transfer costs. (See Appendix A for a different perspective on this point.) In other words, if cash did not exist, investors would hold the asset most like it (the neighboring one on the hierarchy) as the next best solution.

The second nonmonetary aspect of an asset affecting the opportunity cost of money is risk. In an uncertain world, assets have both default risk and risk measured by the variance of its return. Cash has zero default risk and a zero variance in nominal value. Other things equal, risk adverse investors must receive greater returns to compensate for greater risk levels of alternative assets. Thus, the hierarchy of assets with respect to their returns will depend on both transfer costs and risk levels. The opportunity cost of money now equals the return on a zero transfer cost asset of the same riskiness as money.

Implementing the opportunity cost concept requires selecting a zero tranfer cost security with no default risk and a zero variance in nominal value. Some assets have no default risk such as time deposits, various government securities, certain interest bearing checkable deposits, etc. Yet, very few have no variance in nominal value due to penalties on redemption of these assets. Furthermore, the requirement of zero transfer costs narrows the field considerably. From the list of traditional assets, an interest bearing checkable deposit comes closest to matching money's attributes. It has no default risk, minimum variance of return, and zero transfer costs leading many economists to consider it money. Unfortunately, the return on checkable deposits does not measure money's opportunity cost since government price ceilings in this market prevent the return from reflecting buyers' value and sellers' costs.

Recently another candidate has arisen which fills a previous gap in the liquidity hierarchy close to money. Overnight repurchase agreements (RP) have low risk of both default and return variance due to their short maturities. Coupling these properties with very small (mainly fixed) transfer costs has led Tinsley and Garrett (1978) to equate RP with money. In fact, they see the additional use of RP as the reason for the missing money problem of the mid 1970s. If RP effectively represent money, then the market for RP trades money on both sides resulting in a price which represents the true cost of using money. At the margin, this return should approximately equal the total return paid by banks on checkable deposits (explicit plus implicit).

It will only approximate the total return since trading RP does involve some transfer costs. These costs take mainly a fixed form--ticket charge, labor costs, phone bill, etc.--which attracts investors who can spread them over a large purchase. Small investors still prefer checkable deposits even though a part of the return takes an implicit form due to price controls. The return net of all transfer costs for the marginal investor in both markets will be equal. The gross returns of the two assets will also be similar since these assets have the same risk and similar transfer costs. The return on RP gives a superior measure of money's opportunity cost since uncontrolled market forces allow them to accurately reflect buyers' and sellers' desires. Unfortunatley, data on RP returns are not readily available. Yet, a suitable proxy does exist.

RP yields should move with the federal funds rate (FFR) allowing the latter to provide an accurate measure. Since the Federal Reserve system prohibits non-bank traders in the federal funds market, only banks can deal in both RP and federal funds. Their action intergrates the two markets by maintaining a cost based differential between the rate on RP and the FFR. When trading alters this wedge, banks can profit by arbitrage. Such counter buying or selling re-establishes the cost based difference, forcing the two rates to move as one. Thus, the FFR can accurately measure changes in money's price.

Summary

Defining money in terms of transfer costs and risk suggests using a liquidity hierarchy to describe the relationships between assets. Money resides at one end of this spectrum as a riskless medium of exchange. The asset most like cash occupies the neighboring position and has zero transfer costs and the same risk as money. It has a return that measures what cash holders give up to obtain money's attributes. Repurchase agreements represent a useful real world example of such an asset. Since their return moves with the federal funds rate, the latter gives an accurate measure of money's price. Applying the results of this discussion to each model helps create final testable specifications.

Testable Specifications

Inventory Model

Overview

Testable specifications of the inventory model evolve from altering existing functional forms. The resulting adjusted equations necessitate choosing measurable variables. Completing that task renders the final testable forms.

Existing Functional Forms From Previous Work

The Baumol (1952) analysis provides a demand for money equation as,

$$LN(M) = \frac{1}{2} LN(b) + \frac{1}{2} LN(T) - \frac{1}{2} LN(2r)$$

which represents a testable hypothesis. Subsequent work by Tobin (1956), Brunner and Meltzer (1967), and Sprenkle (1966) remove the fixed elasticity predictions. The resulting inventory specification based on deterministic cash flows equals,

$$LN(M) = \beta_0 LN(b) + \beta_1 LN(T) + \beta_2 LN(r).$$

Miller and Orr (1966) derive an inventory variant for money demand as,

$$LN(M) = .192 + \frac{1}{3} LN(b) + \frac{1}{3} LN(\sigma^2) - \frac{1}{3} LN(r).$$

Papers by Miller and Orr (1968) and Frost (1970) make the elasticity values empirical questions. Thus, an inventory variant assuming stochastic cash flows appears as,

$$LN(M) = \alpha_0 + \alpha_1 LN(b) + \alpha_2 LN(\sigma^2) + \alpha_3 LN(r).$$

Adjustments to Existing Functions

In addition to the alterations implied by the discussion of money's price, the literature on cash management techniques suggests the following. (See Appendix B for a detailed discussion.) First, the analysis suggests that the inverse of the volume of repurchase agreement trading gives an accurate measure for transfer costs over the time period under study. Inventory models propose the existence of only one interest bearing asset for firm's portfolios. Thus, transfer costs for the firm equal the cost of moving wealth between this asset and money. For example, if treasury bills represented the only assets firms held, the average transfer costs (ATC) would equal the cost of moving one dollar of wealth between T-bills and money. At first glance, the bid-ask spread might provide a good approximation of such costs since money and T-bills have no risk. The latter property ensures that the asking return will not reflect a risk premium. However, the spread indicates only the ATC of middleman dealing with T-bills. The firm's ATC exceed this amount since submitting a bid involves labor time, phone expenses, etc. Measuring these additional costs is difficult since they become embedded in the bid and, thus, unobservable. The problem increases when dropping the assumption of only one alternative asset.

In reality, a set of assets, along with money, make up a firm's portfolio. Yet, when modeling money demand the relevant ATC equals that on money's best alternative. As relative returns and ATC change, restructuring of the portfolio occurs. To minimize the overall expense of this process, firms transfer wealth between neighboring assets.

Thus, the only transfer cost relevant for money demand is that of checkable and saving deposits at banks. Unfortunately, they remain unobservable on the firm level since they include items like labor costs, phone expenses, etc. A different line of reasoning may provide a practical measure for ATC.

Least squares estimation requires knowledge of the changes in ATC from observation to observation not its exact magnitude for proper estimation. Thus, a variable can serve in the equation if it moves with changes in ATC's exact magnitudes. Such a scaling procedure allows the proxy variable to capture the relationship between ATC and money demand from observation to observation. Porter and Mauskopf (1979) suggest that the growth in repurchase agreement (RP) trading has occurred as a consequence of falling ATC. The latter causes firms to shift portfolio wealth out of money and into interest bearing form. Since RP, out of all assets held, has the lowest transfer costs, firms move wealth from money to RP first and then from RP to longer maturity higher transfer cost assets as a long run decision in response to changes in relative ATC. Thus, the market volume of RP trading has an inverse relationship to the general decline in transfer costs. The existence of data on the latter allow it to provide a useful measure of transfer costs.

Second, Enzler, Johnson, and Paulus (1976) suggest that structural change of the equations should occur with each peak in short term nonbank security rates. These peaks coincide with the acceptance of a new asset into the firm's portfolios. This availability causes a fundamental shift of the relationship between money and the resulting portfolio which can be modeled in the following way.

Capturing structural change necessitates the introduction of dummy variables into the simple inventory equations. The treasury bill rate achieved three peaks during the annual sample time span. Thus, a set of dummy variables must be added to the simple inventory model for each year 1966, 1969, and 1973. For every set, one dummy adjusts the coefficient for transfer costs, another alters the slope for transactions, and a third changes the interest rate parameter.

Third, Porter and Mauskopf (1979) suggest that CMT have lowered both the perceived and actual variances of cash flows. Two ways exist for testing this conjecture with annual firm level data. First, the data permits computation of a cross sectional cash flow variance. In each year, cash flow is constructed for every firm, from which a mean and variance can be formed across all firms. The latter indicates the dispersion of flows over firm size if businesses behave homogeneously. Downward drift in this variance over time would substantiate the CMT conjecture. Second, they argue that sales and cash flow variances had a positive correlation before the introduction of CMT. This relationship allowed sales to serve as a proxy for the unobservable variance. The use of CMT causes the variance to take a route independent of sales. A correlation analysis would validate this claim if a strong correlation exists which deteriorates over time and perhaps becomes negative.

Finally, the use of CMT can be modeled by a stock adjustment process. Firms alter money demand relationships slowly in response to new techniques since incorporating CMT requires time to learn about their unique functions. This process implies the opposite extreme to adding dummy variables which suggests an immediate response. Existing inventory equations can incorporate these alterations once accurate variables are selected.

Selection of Variables

Inventory models postulate the existance of a single interest bearing asset as money's alternative. The opportunity cost of money equals the return available from holding the security. In a world where a hierarchy of assets exists, money's opportunity cost equals the return on its next best alternative regardless of the actual securities contained in a portfolio. As argued previously, the asset most like money, whose return also reflects market forces, is RP. Unfortunately, the return on RP is not readily available on the firm level. However, the federal funds rate (FFR) will move with RP returns allowing the former to serve as an accurate measure.

The transactions variable of inventory models does not lend itself to convenient measurement. In the deterministic variant, transactions represent the predetermined dollar volume of trading conducted by a firm in one period. Thus, it reflects the extent of firms' need for cash balances to conduct operations over the period. Unfortunately, since accounting data does not have an entry for transactions levels a measurable variable must come from available data.

The accounting entries which best reflect the need for money are cash receipts (CR) or cash expenditures (CE) since both move with transactions levels over a given period. The relationship is not exact since payment to a firm may take the form of an increase in a non-cash asset or a decrease in a non-cash liability. Symmetrically, a firm's expenditure could occur with a non-cash asset or a credit of the supplier's asset account. These circumstances constitute a minor part of the totals and, thus, should not affect the basic proportional relationship.

Excluding these minor occurrences, cash receipts over a period equal total sales (S) minus the change in accounts receivable (ΔAR). Cash expenditures equal the cost of goods sold (COGS) plus the change in both inventories (ΔI) and gross property plant and equipment ($\Delta GPPE$), minus the change in accounts payable (ΔAP). Interest payments do not appear in the formula since many firms did not report them. The wholesale price index (WPI) will put these variables into real terms.[2] Two further implications result from the constructon of these flows. First, previous studies have used sales levels to measure the transactions variable. In order to provide direct comparisons, deterministic inventory variants will first use sales and then cash receipts. The a priori expectation being that cash receipts will have a stronger correlation with money demand.

Second, combining cash receipts and cash expenditures gives a firm's cash flow as,

$$CF = S - \Delta AR - \Delta I - COGS - \Delta GPPE + \Delta AP$$

which allows computation of a cash flow mean and variance. This construction implies a way to test the stochastic variant which employs the variance of daily cash flows as the scale variable. While the parameters of such a flow remain unavailable, construction of their quarterly and annual counterparts can occur. Since the variance of both these flows move with the variance of daily cash flows,[3] they provide accurate measures.

---

[2]This index is also referred to as the all commodity price index. A detailed description of all these variables appears in Appendix D.

[3]Miller and Orr (1966) assume that the cash flow due to operations changes t times daily by an amount of either +m or -m with an equal

Data needed to construct these variables in real terms are nominal ones deflated by the WPI. The final step forms testable equations in terms of these proxies.

Final Testable Forms

Deterministic Variants. Four equations representing the deterministic inventory variant will be estimated. The first uses sales (S) as

---

[3](Continued) probability of a positive or negative value. Hence, over an interval of days, the distribution of changes in cash balances will be binomial with a mean equal to

$$E(CF) = ntm(p - q) = 0$$

where p and q represent the probability of a positive and negative change, respectively. The equality of p and q implies that, over n, cash receipts equal cash expenditures and, thus, no drift in cash balances exists. Also, the variance of cash flow changes equals,

$$VAR(CF) = 4ntpqm^2 = ntm^2.$$

This distribution approaches normality as n and t increase and implies a corresponding distribution for levels of cash flow.

The cash flow in any period of length $\frac{1}{t}$ equals

$$CF_t = m$$

where m can take a positive or negative value. m has

$$E(m) = 0$$

$$VAR(m) = E(m_i m_j) = \sigma^2 \text{ if } i = j$$

$$= 0 \text{ if } i \neq j.$$

The quarterly cash flow equals

$$CF_Q = \sum_{i=1}^{T} m_i$$

where i sums over all 1/t periods each day for 90 days. The expected value and variance of CF_Q equal

$$E(CF_Q) = E(\sum_i m_i) = 0$$

the transactions variable, the federal funds rate (FFR) as money's price and omits the inverse of repurchase agreement volume (RPV)--the measure of transfer costs. This specification provides direct comparison with previous studies and appears as

$$LN(M) = \beta_0 + \beta_1 LN(S) + \beta_2 LN(FFR) + \varepsilon. \qquad (3\text{-}1)$$

The second equation includes RPV and provides contrast to previous estimates:

$$LN(M) = \beta_0 LN(RPV) + \beta_1 LN(S) + \beta_2 LN(FFR) + \varepsilon. \qquad (3\text{-}2)$$

The third specification employs the alternative transactions variable--cash receipts (CR)--while omitting RPV. Again, this simplification allows for direct comparisons with previous studies.

$$LN(M) = \beta_0 + \beta_1 LN(CR) + \beta_2 LN(FFR) + \varepsilon. \qquad (3\text{-}3)$$

Finally, the last form has CR and RPV as relevant variables.

$$LN(M) = \beta_0 LN(RPV) + \beta_1 LN(CR) + \beta_2 LN(FFR) + \varepsilon \qquad (3\text{-}4)$$

The Baumol-Tobin analysis makes the following predictions. All coefficient estimates should have magnitudes significantly different

---

[3](Continued)

$$\begin{aligned} VAR(CF_Q) &= E(\sum_i m_i)^2 - [E(\sum_i m_i)]^2 \\ &= E(m_1 + m_2 + \ldots + m_t)^2 \\ &= T\sigma^2. \end{aligned}$$

Thus, the variance of quarterly cash flows moves with the variance of the distribution generating cash flows. This relationship means that the variance of CF_Q gives an accurate proxy for daily cash flow variance. The same analysis applies to the annual sample.

from zero. β_1 and β_2 should equal one-half and a negative one-half, respectively. For Equations (3-1) and (3-3), the magnitude of β_0 is not predicted. In Equations (3-2) and (3-4), β_0 should equal one-half.

The literature on cash management techniques implies the following alterations of Equation (3-4). For each previous peak in the treasury bill rate (TBR), a new asset has emerged causing a shift in the parameters of the model. To capture this effect, a dummy variable will be added for 1966, 1969, and 1973 corresponding to peaks in TBR. The model becomes,

$$\begin{aligned} LN(M) = {} & \beta_0\, LN(RPV) + \beta_1\, LN(CR) + \beta_2\, LN(FFR) + D_{66}\, \beta_0\, LN(RPV) \\ & + D_{66}\, \beta_1\, LN(CR) + D_{66}\, \beta_2\, LN(FFR) + D_{69}\, \beta_0\, LN(RPV) \\ & + D_{69}\, \beta_1\, LN(CR) + D_{69}\, \beta_2\, LN(FFR) + D_{73}\, \beta_0\, LN(RPV) \\ & + D_{73}\, \beta_1\, LN(CR) + D_{73}\, \beta_2\, LN(FFR) + \varepsilon. \end{aligned} \tag{3-5}$$

The analysis predicts that the coefficient estimates on CR, RPV, and FFR slope dummies should take significant values. This result occurs since firms want less money for a given amount of these varibles when a new asset is available.

Finally, the incorporation of CMT may require a time lag as firms learn about these methods. A stock adjustment model can capture this effect as,

$$LN(M) = \beta_0\, LN(RPV) + \beta_1\, LN(CR) + \beta_2\, LN(FFR) + \beta_3\, LN(M_{-1}) + \varepsilon. \tag{3-6}$$

The parameters will take values as predicted above. In addition, β_3 should differ significantly from zero.

Stochastic Variants. Two equations representing the stochastic variant of an inventory model will be estimated. The first omits RPV in order to provide direct comparisons with other studies. It also uses cash flow variance (V) as the transactions variable and the FFR as a measure of money's opportunity cost.

$$LN(M) = \alpha_0' + \alpha_1 \, LN(V) + \alpha_2 \, LN(FFR) + \varepsilon \, . \qquad (3\text{-}7)$$

The second equation adds RPV to the specification above,

$$LN(M) = \alpha_0' + \alpha_0 \, LN(RPV) + \alpha_1 \, LN(V) + \alpha_2 \, LN(FFR) + \varepsilon \, . \qquad (3\text{-}8)$$

The Miller-Orr analysis predicts the following results. All coefficients should differ significantly from zero. For both equations, α_0 and α_1 should equal one-third while α_2 should take a value of negative one-third.

Summary

These equations embody the essence of inventory theory. Money demand depends on an interest rate and transfer costs and exhibits economies of scale in transactions. Deriving the specifications required adjusting existing functional forms and then selecting measurable variables. To ensure comparability of inventory specifications with those from other theories, this procedure must branch out to encompass the portfolio model.

Portfolio Model

Overview

Deriving a testable specification of the portfolio approach requires the following process. First, the existing functional form from previous work is presented. The second step adjusts it to fit into a comparable format. Third, variables are selected to accurately represent the theoretical concepts. These steps culminate in a final testable form.

Existing Functional Form From Previous Work

The portfolio approach of Friedman (1956) postulates real money demand as,

$$M = f(w, r_b, r_e, r_p, U)$$

where w represents real wealth, r_i stands for the return on bonds, equities, and physical assets, respectively. Finally, U symbolizes the tastes and preference of the holder.

Adjustments to Existing Functional Forms

One striking difference between Friedman's portfolio equation and most demand functions is the absence of own price. As argued previously, the opportunity cost of money cannot equal the returns on substitute assets included in the function. Those returns measure the value of alternative stores of wealth and influence money balances independently from the cost of holding money. In geometric terms, they cause a shift of the demand curve not a movement along the curve. The latter adjustment occurs due to changes in the return of a zero risk transfer costless asset (r_m), which should appear in the function. Also, firms only invest in commodities which aid in their production process. Since changes in the gross national product price deflator measure the return on all physical goods, they do not accurately measure the return for an individual firm's investment. A product price index for specific physical goods purchased by each firm provides a more accurate variable. The next section selects the remaining variables necessary to create final testable forms.

Selection of Variables

The portfolio model describes firms' money demand as a function of real wealth, the opportunity cost of money, the rate on interest earning securities, the return on equities, the return on physical assets, and tastes. These will be explored in turn.

Nominal current assets deflated by the wholesale price index (CA) can measure the wealth constraint. For an individual, savings becomes an addition to wealth. Symmetrically, the retained earnings of a firm add to current assets. Thus, the latter provides a direct measure of a firm's wealth. Previous study's use of sales instead of wealth does not capture the unique aspect of a portfolio approach. Furthermore, its use is unnecessary since a direct measure of wealth exists.

Selecting rates of return must occur in light of the reason portfolios contain many assets. Firms diversify wealth in order to gain the variety of benefits available from a spectrum of assets. Money balances render perfect liquidity and zero risk. The price of holding cash equals the rate on federal funds (FFR)--a measure of the rate on money's closest alternative. Capturing the diversity property necessitates selecting an interest bearing security with the opposite characteristics of money. Since the ten year corporate bond rate (CBR)[4] is a long maturity asset with both high transfer costs and high risk, it will serve to measure r_b.

Equity returns should reflect the firm's best opportunity to participate in the ownership rights of some business. The highest valued opportunity for most firms is re-investment in themselves.[5] Thus, the firm's own equity return (EQR) will measure r_e.

Finally, the return on physical assets equals the capital gain accruing over time due to appreciation. The many commodities held by firms in the course of business will be products of other firms. Thus, for each firm the percentage change in the product price index (PPI) in the industry supplying them capital goods will measure the rate on physical assets. These variables help form a testable equation of the portfolio approach.

Final Testable Forms

Overview. The exact specification for the portfolio model is revealed by employing the Box-Cox analysis. Since this method selects a form on the basis of the sample, it must be used on both the quarterly and the annual data. Completing this task reduces the chance of specification error.

Friedman Variant. For both samples, the Box-Cox analysis suggests log linear as the optimal form. The portfolio equation appears as,

---

[4]The AA bond rate was selected to reflect a higher risk level.

[5]This point suggests that the firms' owners have found investment in their business as preferable to alternatives.

$$LN(M) = \beta_0 + \beta_1 LN(CA) + \beta_2 LN(FFR) + \beta_3 LN(CBR) + \beta_4 LN(EQR) + \beta_5 LN(PPI) + \varepsilon \qquad (3\text{-}9)$$

where the regressors are real current assets (CA), the federal funds rate (FFR), the ten year corporate bond rate (CBR), equities returns (EQR), and the percentage change in the product price index (PPI). The model predicts that β_1 should exceed zero while all rates of return parameters should take negative values. Finally, portfolio theory makes no predictions about parameter magnitudes.

Summary

The Friedman equation, appropriately amended, embodies the essence of the portfolio approach as it applies to firms. Wealth appears as the scale factor and acts as a constraint on holding additional money. In contrast, inventory models view the scale factor as measuring the need for money. They include transactions and justify it on the idea that making a dollar's worth of payments requires a fixed stock of money. The portfolio approach treats this relationship as a result of the utility maximum process (Friedman, 1956). If the cost of holding money increases, then it pays to reduce the volume of transactions per dollar of cash balances. The money demand function should include these basic conditions causing holders to substitute away from or into money. Since the various rates of return determine the optimum relationship between money balances and transactions levels, no need exists to include the latter. Incorporating these characteristics into a specification required altering existing functional forms and selecting measurable variables. Finally, this procedure must extend to envelop the production model.

Production Model

Overview

Previous work provides three functional forms embodying the production approach. After adjusting them to fit into comparable form, variables are selected that accurately measure the relevant theoretical concepts. This procedure renders three final specifications.

Existing Functional Forms From Previous Work

Saving (1972) derives a transactions cost variant for money demand as,

$$M = g(P,\ H_X,\ H_V,\ H_M)$$

where P represents the price of output and H_X, H_V, and H_M stand for the opportunity cost of holding output, physical inputs, and money, respectively. Nadiri (1969) forms a neoclassical variant,

$$M = f(Q,\ r_m, r_k, r_l)$$

as firms' money demand. Q represents output levels, and r_m, r_k, and r_l the opportunity cost of using money, capital, and labor, respectively. These two equations provide a basis for testable specifications. As an alternative, Dennis and Smith (1978) assume a specific production process that leads directly to a testable form. The translog cost variant appears as,

$$\begin{bmatrix} C_l \\ \\ C_m \end{bmatrix} = \begin{bmatrix} 1 & 0 & LN\frac{P_l}{P_k} & LN\frac{P_m}{P_k} & 0 & LN(Q) & 0 \\ 0 & 1 & 0 & LN\frac{P_l}{P_k} & LN\frac{P_m}{P_k} & 0 & LN(Q) \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \gamma_{11} \\ \gamma_{21} \\ \gamma_{22} \\ \delta_1 \\ \delta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}$$

where C_l and C_m represent the relative cost shares of labor and money, respectively. P_i stands for the price of the i^{th} input and Q is real output.

Adjustments to Existing Functional Forms

Previous discussions suggest that the federal funds rate represents an accurate measure of the price of money. The three production models need no further adjustments since they are formed specifically to describe firm behavior. The final task selects the remaining measurable variables.

Selection of Variables

Each production equation postulates that all input prices affect money demand. In addition, the transactions cost variant adds both the price and the holding cost of output to explain cash balances. Both the neoclassical and translog variants use the amounts of output as regressors with input prices. Finally, the translog equations also require the construction of relative cost shares. Selecting measurable variables will be discussed for each concept in turn.

Beginning with input prices, the opportunity cost of money equals the rate on federal funds (FFR). The price of labor equals wage rates for specific census code commodities (WR). Identifying each firm's major commodity allows this data to accurately measure the price of labor. Finally, construction of the opportunity cost of capital (RR) follows the method of Hall and Jorgenson (1967). The measure accounts for both the price of new capital goods and the discounted value of their future services.

Turning to output, the price charged by firms is a price index figure available for specific commodities (PPC). Matching the firm with its major commodity lets this data properly reflect output's price. The opportunity cost of holding output equals the firm's rate of return on equity (EQR). At the margin, this return will be identical to the rate on the firm's best alternative. Finally, physical output levels (Q) equal dollar sales plus the change in inventory, divided by the relevant product price index for a specific commodity.

The relative cost share for the i^{th} input (C_i) takes the dollar amount spent on i and divides it by the total expenditure on all inputs. The expenditure on labor comes directly from financial statement entries. The total amount given up to use money is computed as price times quantity. Capital's cost follows the same calculation, which requires constructing the real amount of capital used. The latter equals the change in capital stock deflated by the price of capital goods. Given these measures, the next step writes explicit testable equations.

Final Testable Forms

Overview. Finding the exact form for the transactions cost and neoclassical variants requires using the Box-Cox analysis. The method must be applied to both the annual and quarterly samples since it selects a form based on the data. The results enhance the probability of correctly specified equations.

Transactions Cost Variant. The Box-Cox analysis suggests a linear in natural logarithm form for this variant. Specifically, the transactions cost equation appears as,

$$LN(M) = \beta_0 + \beta_1 LN(PPC) + \beta_2 LN(FFR) + \beta_3 LN(EQR) + \beta_4 LN(WR)\ \beta_5 LN(RR) + \varepsilon \tag{3-10}$$

where PPC represents the price of output, while FFR, EQR, WR, and RR stand for the opportunity cost of money, output, labor, and capital, respectively. β_1 should exceed zero while β_2 and β_3 should be less than zero. β_4 and β_5 can take either sign. This variant makes no predictions about parameter magnitudes.

Neoclassical Variant. The Box-Cox analysis indicates a log linear form, making this variant,

$$LN(M) = \beta_0 + \beta_1 LN(Q) + \beta_2 LN(FFR) + \beta_3 LN(WR) + \beta_4 LN(RR) + \varepsilon \tag{3-11}$$

where Q stands for output levels, while FFR, WR, and RR represent the price of money, labor, and capital, respectively. This variant predicts a positive value for β_1 and a negative value for β_2. β_3 and β_4 can take either sign. No parameter magnitudes are predicted by the model.

Translog Cost Variant. This variant has a specific form equal to,

$$\begin{bmatrix} C_1 \\ \\ C_m \end{bmatrix} = \begin{bmatrix} 1 & 0 & LN\frac{WR}{RR} & LN\frac{FFR}{RR} & & 0 & LN(Q) & 0 \\ \\ 0 & 1 & 0 & LN\frac{WR}{RR} & LN\frac{FFR}{RR} & 0 & & LN(Q) \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \gamma_{11} \\ \gamma_{21} \\ \gamma_{22} \\ \delta_1 \\ \delta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}$$

where C_i represents the relative cost share of the i^{th} input an Q represents the level of output. All parameters should differ significantly from zero and allow construction of the production and input demand structures.

Summary. The production model renders a money demand function dependent on the underlying functions of the production process. Thus, justification exists for forming a transaction cost, neoclassical, and translog cost versions. In each case, money demand is derived from the maximization process and depends on the level and prices of output and inputs. To ensure comparability, the specifications were formed by changing existing functions. The resulting adjusted equations, when coupled with measurable variables, become final testable forms.

Summary

This section has developed 12 specifications of firms' money demand. In each case, the following adjustments applied to existing functions, given by previous studies. First, all equations were put into real terms ensuring that each one describes the same dependent variable. Second, the federal funds rate gives an accurate measure for the opportunity cost of money. After selecting the remaining variables, final testable forms emerge with two characteristics. They embody the unique aspects of their corresponding theory and provide for accurate comparisons of performance. Those evaluations employ statistics generated by combining the final two parts of the methodology.

Econometric Techniques

Overview

Ordinary least squares gives the fundamental econometric method used in the dissertation. This selection is based on the properties of unbiasedness and efficiency of the estimators which result under classical linear regression assumptions. Four sets of circumstances arise in this study which violate these assumptions and, thus, require more advanced methods.

First, omission of a relevant variable and/or pooling data requires an ajustment by either the covariance or error components models. Second, general functions must take a specific form before estimation can occur. The Box-Cox analysis fulfills this task in an optimal way. Third, multicollinearity may exist in equations with correlated independent variables. Ridge regression locates and cures this problem. Finally, sets of equations require a systems method of estimation. Both two-stage Zellner and three-stage least square represent such techniques. These methods will be outlined in turn.

Pooling Adjusted Methods

Covariance Model

When pooling data from a number of firms across time, the probability runs high that parameters differ both between firms and between time periods. The covariance model, through the introduction of dummy variables, can partially account for such divergences. Alternatively, the method can be viewed as an adjustment for omitting a relevant variable whose value differs among firms and across time. Thus, the covariance model provides a way to account for transfer costs (in inventory models) or tastes (in portfolio models) which differ for each unit and time period.

The procedure adds a dummy variable, to adjust the basic model's intercept, for all but one firm and for all but one period. The covariance model, for the simple regression of X on y, appears as,

$$y_{it} = \alpha + \beta X_{it} + \gamma_2 w_{2t} + \ldots + \gamma_N w_{NT} + \delta_2 z_{i2} + \ldots + \delta_T z_{iT} + \varepsilon_{it}$$

where w_{it} equals one for the i^{th} firm and zero elsewhere while z_{it} is one for the t^{th} period and zero otherwise. The model contains (N-1) + (T-1) dummy variables, two less than the number of firms plus periods to prevent perfect collinearity.

If the disturbance satisfies the classical linear regression model assumptions, OLS will render unbiased and consistent estimates of all parameters. Those for the dummy coefficients measure the change in the intercept term's value with respect to the first firm and first time period. In other words, α captures the effect of γ_1 plus δ_1. Thus, each of the remaining dummy coefficients measure deviations from α--the true intercept.

The choice between the simple pooled model and the covariance model, which sacrifices degrees of freedom, is properly made using a statistical test. The test weighs the significance of the joint explanatory power of all dummy variables. A statistic, distributed f under the null hypothesis, measures the additional explanatory power from adding dummies compared to the total explanatory power of the covariance model. If the dummies have joint significance, then they must appear in the equation. This inclusion allows OLS to give unbiased estimates of the remaining coefficient parameters.

The inclusion of dummy variables into the simple pooled model adjusts systematically for lack of knowledge about the equation. The omitted variables are assumed to follow a pattern, changing for every firm and period. An alternative technique models this ignorance differently.

Error Components Model

The relationship between the covariance model and the error components model stems from their views about the intercept term. The covariance model sees it as a set of fixed parameters. The error components approach models it as two random variables, one arising from cross-sections the other from time series. If both random terms have a normal distribution, the error components model saves degrees of freedom relative to the covariance model since estimation with the latter requires knowledge of only a mean and a variance of each component as opposed to a set of parameters. The error components model assumes the mean effect of each random term gets captured by the model's intercept term. The random deviation about these means become components of the model's disturbance term. Formally, the error term consists of three components as,

$$\varepsilon_{it} = U_i + V_t + W_{it} \qquad i = 1 \ldots N$$
$$t = 1 \ldots T$$

where U_i represents the cross-section component, V_t stands for the time series component, and W_{it} is a combined effect. For each part of ε_{it},

$$U_i \sim n(0, \sigma_U^2)$$

$$V_i \sim n(0, \sigma_V^2)$$

$$W_{it} \sim n(0, \sigma_W^2)$$

which states that each component has a normal distribution with zero mean and constant variance. Furthermore,

$$E(U_i V_t) = E(U_i W_{it}) = E(V_t W_{it}) = 0$$

$$E(U_i U_j) = 0$$

$$E(V_t V_s) = 0$$

$$E(W_{it} W_{is}) = E(W_{it} W_{jt}) = 0$$

which states that the various components have no cross correlations either between or with themselves.

These properties imply the following.[6] First, the model has a homoscedastic variance equal to,

$$VAR(\varepsilon_{it}) = \sigma^2 = \sigma_U^2 + \sigma_V^2 + \sigma_W^2 \,.$$

Second, the coefficient of correlation between the disturbances of two cross-sections at a given time is a constant. Third, the coefficient of correlation between disturbances in two time periods for a given cross-section remains fixed over time. This pattern contrasts to a disturbance with first order autocorrelation where the correlation between this

---

[6]These properties are fully outlined by Maddala (1971).

period's disturbance and a past disturbance declines as time regresses. Finally, the coefficient of correlation between different cross-sections in different periods equals zero.

Combining these results implies a covariance matrix with positive non-diagnonal elements. Under these conditions, obtaining unbiased and efficient estimators requires using Aitken's generalized least squares procedure. Unfortunately, the elements of the covariance matrix are unknown parameters. Yet, an estimator retaining the same asymptotic properties as Aitken's can be constructed by employing a consistent estimate of the covariance matrix in the above equation. Such an estimate occurs by using the residuals from OLS on the simple pooled model to construct consistent estimates of the components' variances (Maddala, 1971). A consistent and asymptotically efficient estimator of the regression parameters results by substituting the covariance matrix estimates into Aitken's generalized least squares formula. This estimator can be more efficient than the covariance model estimator when the data has a disturbance with the properties assumed by the error components model. Specifically, the latter has a fixed correlation over time which contrasts with a first order autocorrelation scheme. Tests can reveal the actual generating process and, thus, can indicate whether the error components model gives more efficient estimators.

Selecting a Specific Form

The Box-Cox technique provides a method of finding an equation's optimal functional form (Box and Cox, 1964). A maximum likelihood criteria serves as the standard for an optimum. The technique essentially selects the best exponential transformation of the variables in terms of the criteria. One advantage of selecting the power function as a base form is that it includes the simple linear regression as a special case--a degree one power function. Thus, linearity can be tested against alternative forms of a higher degree. The specific function appears as,

$$\frac{Y_i^\lambda - 1}{\lambda} = \alpha + \beta \left(\frac{X_i^\lambda - 1}{\lambda}\right) + \varepsilon_i.$$

For each value of λ, the function raises variables to different power. For example, when λ equals one, the expression becomes, a simple linear regression model. More generally, for any λ greater than zero, the equation regresses Y_i on X_i each raised to the λ power. For λ equal to a negative one, the equation reduces to a regression of the inverse of the independent variable on the inverse of the dependent varible. In general, for any λ less than zero, the technique regresses the reciprocal of Y_i on the reciprocal of X_i each raised to the absolute value of λ power. If λ equals zero, the expressions for the dependent and explanatory variable equal LN(Y) and LN(X), respectively. The equation, for λ equal to zero regresses the natural logarithm of X on similarly transformed Y values.

By selecting different values for λ various regression forms can be estimated and tests conducted to select the optimum form. To carry out the test, estimates of λ, its standard error, and the corresponding regression parameters can be generated by the maximum likelihood method. For each λ, the corresponding parameter estimates which create the maximum value for the likelihood function can be found along with standard errors.

The optimal form of the equation is defined as the transformation resulting in the highest value of the likelihood function given all λ values. Implementing this criteria necessitates that for each selected λ value variables are transformed, a constrained regression run, and a likelihood function computed. The form, determined by the value of λ, generating the largest value for the likelihood function, given maximum likelihood estimates, is the optimum transformation. This procedure also suggests a way to test the form.

For example, the test for a log linear relation has the structure,

$$H_0: \quad \lambda = 0$$

$$H_1: \quad \lambda \neq 0$$

and a test statistic formed as $\chi^2 = 2(L_M - L_0)$. Where L_M stands for the value of the likelihood function at λ's optimal value and L_0 represents the function's value when λ equals zero. The statistic has, asymptotically, a Chi-square distribution with one degree of freedom. H_0 would be rejected if the calculated statistics exceeded the table value.

Multicollinearity

The third econometric problem--multicollinearity--occurs when the $X'X$ matrix of the OLS estimator has a determinant near zero. This property results when data for two or more explanatory variables change in a nearly fixed proportion. Such linear combinations of economic data are common, especially with respect to rates of return. Ridge regression gives a diagnostic tool to locate the related variables and offers estimators superior to OLS. Locating the offending variables requires exploiting the imprecision of OLS estimates while finding superior estimators traces their instability. Both solutions occur in the following framework.

Essentially, ridge regression inflates the diagonal elements of the $X'X$ matrix by a factor k. Because this transformation increases the determinant of $X'X$, it eliminates the near singularity and reduces the diagonal elements of $(X'X)^{-1}$. The latter implies that the resulting coefficient estimators have both smaller variances and values biased toward zero. If the variance declines relative to the bias squared, the estimators formed using k have lower mean square errors than OLS estimators. With a smaller mean square error, an estimator will, on average, give estimates lying closer to the true parameter value.

To perform the diagnostic function, regressions with successively larger k values are estimated. This process "generates" additional information by altering the $X'X$ used to compute estimates. Related explanatory variables should have coefficient estimates sensitive to the addition. These unstable coefficients deflate toward zero as k increases and belong to the related explanatory variables. In addition to diagnosing the source of the problem, increasing k implies the following.

If deflation occurs, the minimum mean square error estimators exist approximately when these coefficients stabilize. As shown by Hoerl and Kennard (1970), at that point, the following properties hold. The system acts as an orthogonal one, implying the absence of multicollinearity. The residual sum of squares shows no unreasonable enlargement relative to the minimum residual sum of squares--from OLS. Finally, coefficient estimates will change to the hypothesized signs and magnitudes. Thus, ridge regression provides a useful method to identify and deal with multicollinearity.

Systems Methods

The last group of techniques deals with equation sets. Estimation of the translog cost function presents the problem of seemingly unrelated regressions. Correlation exists between the disturbance terms of the cost share equations because errors in carrying out ideal profit maximum outcomes occur simultaneously for all resources. Under this condition, OLS remains consistent but not asymptotically efficient. A two-stage Zellner (2SZ) technique provides more efficient estimators by accounting for the cross equation correlation (Zellner, 1962). The first stage estimates each equation in the set by OLS. The resulting consistent estimates form a consistent estimate of the disturbance term's covariance matrix. The second stage incorporates that matrix in a generalized least squares process which gives estimators with the same asymptotic properties as Aitken's GLS estimators (Dhrymes, 1970). This equivalence implies that the 2SZ estimates are consistent and asymptotically efficient. However, they are also sensitive to selection of the omitted equation.

Overcoming this difficulty requires iteration of the two-stage process until the estimates converge. Such a repetition makes the 2SZ estimates identical to maximum likelihood estimates. Therefore, they will be consistent and asymptotically efficient regardless of which equation is omitted.

Finally, if the only source of cross equation correlation is the disturbance term, the 2SZ estimators are both consistent and asymptotically efficient. The most common violation of this condition occurs when the dependent variable of one equation serves as an explanatory variable in another equation. The translog set does not contain such correlation. Yet, another source may present a problem.

Violation can also occur if stochastic variables, interdependent with the disturbance terms, serve as explanatory variables. The prices of inputs pose no threat since competitive individual firms take them as given. However, output levels may exhibit a stochastic relationship with the disturbance terms. This correlation could arise due to errors in the profit maximum process which causes not only the disturbance terms, but output deviations as well. In this case, 2SZ estimators remain consistent but not asymptotically efficient. The efficiency loss occurs because the cost share equations now make up a truly simultaneous system and, thus, require a more powerful method of estimation.

Three stage least squares (3SLS) provides such a technique (Zellner and Theil, 1962). The increase in efficiency results from incorporating the cross equation correlation in three stages. In the first step, OLS generates consistent estimates of the independent variables. These estimates form consistent fitted values for the independent variable that have no correlation with the disturbance terms of other equations. They serve as instruments in the second stage. There, the structural equations are estimated individually by substituting the instruments for explanatory stochastic variables. The results can calculate a consistent estimate of the disturbance term's covariance matrix which will contain estimates of the cross equation correlation. The final stage incorporates the covariance matrix in an Aitken's generalized least squares method. This technique gives estimators which are both consistent and asymptotically efficient. Yet, they are sensitive to the choice of the omitted equation.

Iteration of the three stages solves this problem since the estimators converge to maximum likelihood estimators (Dhrymes, 1970). The latter are consistent and asymptotically efficient regardless of the omitted equation. Iterative versions of both 3SLS and 2SZ estimators will be used on the translog model.

Summary

Six advanced econometric techniques are required to implement an econometric analysis of firms' money demand. Either the covariance model or the error components model can compensate when a specification omits a relevant variable. Transfer cost in inventory equations and tastes in the portfolio model represent such variables. Alternately, the two methods can be viewed as adjustments for pooled data and, thus, apply to all three theories. The Box-Cox method employs a maximum likelihood criteria to select specific forms of the portfolio, transactions cost, and neoclassical equations. For these same three, ridge regression indicates the presence of multicollinearity and provides alternative estimators to ordinary least squares. Two-stage Zellner and three-stage least squares generate efficient estimates of the translog cost system of equations. Implementing these techniques to measure each theory's performance, requires the third part of the methodology.

Samples

Overview

The samples used in previous money demand studies did not provide a basis for testing between competing models. Overcoming this inadequacy requires a sample to have two unique characteristics. First, it must contain observations on all variables postulated by all equations. This completeness allows estimation of each model over the same set of circumstances and, thus, aids in judging relative performance. Second, the sample must contain firm level data since the models postulate individual firm behavior. Relying on aggregated data to test microrelations requires stringent assumptions. Validating these conditions for a data base necessitates using firm level observations. If no aggregation bias exists, then testing can employ aggregate data. The presence of such bias in large amounts casts doubts on any conclusions made from aggregated data. A central feature of this dissertation is the use of individual firm observations for testing firm's money demand.

Certain advantages exist in using both quarterly and annual data bases. Thus, two samples of panel data will be used. Both contain firm level observations for all variables over their respective time periods.

Quarterly Sample

One takes quarterly measurements from 1972-4 through 1980-1 resulting in 30 observations per firm. The advantage of such short run data pertains to matching flows and stocks. For example, the deterministic inventory model relates the level of cash holdings (stock) to the rate of transactions (flow). Consider measuring this relationship with annual data. Money balances on December 31 would be linked with a 12-month accumulation of transactions. If a firm had strong business for 11 months and then slumped badly in the 12^{th} month, the annual accumulation of transactions would reflect the prosperous period while cash balances would correspond to the slump. Even though quarterly data may be subject to the same problem, it should suffer to a smaller degree.

The sample was constructed as follows.

Beginning with a list of Standard and Poor's Top 500, firms were deleted on the basis of incomplete data. The resulting quarterly sample contains 95 firms ranging in size from annual sales of one million dollars up to the largest corporations in the United States. Nineteen industries are represented covering the spectrum from heavy manufacturing and mining to retail trade to services. The only available sector deliberately omitted was public utilities due to the weakness of the profit motive in such areas. Each theory bases money demand on the postulate of profit maximizing behavior. Without such a stimulus, firms may not act in accordance with the model's predictions. The only other reason for an industry not appearing in the sample was incomplete data of all its firms. Data availability also dictated the time span of the sample. Thirty observations per firm exist for 1972-4 through 1980-1.

Annual Sample

The other sample lists annual figures from 1962 through 1979 which gives 18 observations per firm. Annual data have two major advantages: They allow tests of certain hypotheses which cannot be conducted over a short period of time and they contain a broader spectrum of variables. The latter condition allows estimation of additional forms of the models which contain variables not measured on a quarterly basis.

The annual sample follows the same construction as the quarterly one. One hundred ninety-two firms survived the complete data test for 1962 through 1979. The same size range exists as for the quarterly sample. The firms represent 28 industries; as before, public utilities were purposely excluded. The variables and sources which determined the two samples appear in Appendix D.

Summary

The two samples contain facts that allow a logical sequence in testing money demand functions. This conclusion stems from their two distinguishing characteristics; individual firm observations and complete data on every variable postulated by all models. Combining these samples with the econometric techniques discussed previously provides statistics used in model evaluations. As a prerequisite, the next section discusses finding the correct plane for forming the estimates.

Aggregation Analysis

Theil (1971) outlines a method of constructing the aggregation bias contained in each macroparameter estimate. The procedure computes the expected value of any macroparameter from estimates of microrelations, auxiliary regressions and a macrorelation. To illustrate, the deterministic inventory model has money demand for the i^{th} firm from a group of N firms as,

$$LN(m_i) = \beta_{0i} + \beta_{1i}\, LN(t_i) + \beta_{2i}\, LN(r_i) + \varepsilon_i$$

where i = 1 . . . N and the β_{hi} stand for microparameters where h = 0, 1, 2. m, t, and r represent the i^{th} firm's money holdings, transactions amounts, and interest rate, respectively. These three are microvariables of the above microrelation. Macrovariables for money holdings, transactions levels, and the interest rate result from geometric summation of the microvariables over N firms as,

$$M = [m_1\, m_2 \ldots m_N]^{1/N} \quad T = [t_1\, t_2 \ldots t_N]^{1/N} \quad R = [r_1\, r_2 \ldots r_N]^{1/N}.$$

This process means that,

$$LN(M) = \frac{1}{N} \sum_i LN(m_i)$$

which renders the correct macrorelation by substituting for $LN(m_i)$ to form,

$$LN(M) = \frac{1}{N} \sum_i \left[\beta_{0i} + \beta_{1i}\ LN(t_i) + \beta_{2i}\ LN(r_i) + \varepsilon_i\right]$$

$$LN(M) = \frac{1}{N} \sum_i \beta_{0i} + \frac{1}{N} \sum_i \beta_{1i}\ LN(T) + \frac{1}{N} \sum_i \beta_{2i}\ LN(R) + \frac{1}{N} \sum_i \varepsilon_i .$$

However, this last expression is not a linear equation in the macrovariables. Yet, estimation using macrovaribles employs a linear equation as,

$$LN(M) = \beta_0 + \beta_1\ LN(T) + \beta_2\ LN(R) + \varepsilon_i$$

which is called an incorrect macrorelation. Here the β_h stand for macroparameters, where h = 0, 1, 2. Their estimates contain aggregation bias since the incorrect macrorelation is a misspecified version of the correct macrorelation.

Finally, the auxiliary regressions for the i^{th} firm, which illuminate the misspecification contained in β_1 and β_2, respectively, take the form,

$$\frac{1}{N}\ LN(t_i) = \alpha_{1i} + b_{11i}\ LN(T) + b_{12i}\ LN(R) + \varepsilon_{1i}$$

$$\frac{1}{N}\,LN(r_i) = \alpha_{2i} + b_{21i}\,LN(T) + b_{22i}\,LN(R) + \varepsilon_{2i}$$

where the b_{hki} represent regressor coefficients and $h = 1, 2$ and $k = 1, 2$. Estimation of the N auxiliary regressions with $\frac{1}{N}\,LN(t_i)$ as a dependent variable, the incorrect macrorelation, and the N microrelations provide the estimates needed to construct the expected value of the macroparameter estimate for T. This expectation would equal

$$E(\hat{\beta}_1) = \frac{1}{N}\sum_i \beta_{1i} + \sum_i(b_{11i} - \frac{1}{N})\beta_{1i} + \sum_i b_{12i}\,\beta_{2i}$$

where β_{1i} and β_{2i} represent true (unknown) microparameter values. To allow construction, they will be assumed equal to their estimates from the microrelations. The expected value of the macroparameter for R requires estimates from the N auxiliary regressions using $\frac{1}{N}\,LN(r_i)$ as a dependent variable, the incorrect macrorelation, and the N microrelations. In general, the expected value of the macroparameter of the h independent macrovariable equals,

$$E(\hat{\beta}_h) = \frac{1}{N}\sum_i \beta_{hi} + \sum_i(b_{hhi} - \frac{1}{N})\beta_{hi} + \sum_{k\neq h}\left(\sum_i b_{hki}\,\beta_{ki}\right)$$

whose construction necessitates using only the N auxiliary regressions with the h microvariable as an independent variable, the N microrelations, and the incorrect macrorelation.

Once constructed, each component of this expectation has the following meaning. The first right hand term represents the true value of the macroparameter as given by the correct macrorelation. The last two righthand terms constitute the total bias due to aggregation since they cause the expected value of the macroparameter estimate to differ from its true value. The second righthand term generates bias from corresponding microparameters while noncorresponding microparameters cause the bias given by the final term.

Construction of the components of $E(\hat{\beta}_h)$ for each model occurred as follows. First, the macrovariables were formed by geometric averaging of the firm level data. OLS estimation of the resulting incorrect macrorelation gives the $\hat{\beta}_h$. Second, OLS estimated the 95 microrelations each using the data of one firm to render the $\hat{\beta}_{hi}$ and $\hat{\beta}_{ki}$. Third, the auxiliary regressions underwent OLS estimation resulting in the b_{hhi} and b_{hki}. Finally, the formula for $E(\hat{\beta}_h)$ was applied using these estimates. Once constructed, the bias can be tested for significance.

The test for aggregation bias comes from Zellner (1962) and has a base in the general f test for constrained and unconstrained regressions. Since the general f test applies in many cases throughout this chapter, an explanation of the procedure appears next. Then its use in the specific case of aggregation bias will be outlined.

Given a testable restriction contained in a null hypotheses, a constrained equation is estimated assuming the validity of the restriction. The resulting estimates construct an error sum squares (ESS_R) which measures the unexplained portion of the variation in the dependent variable. Then an unconstrained equation undergoes estimation which postulates invalid restrictions. The resulting error sum squares (ESS_U) indicates the inaccuracy of using the equation to estimate dependent variable values. A test statistic appears as,

$$f = \frac{\dfrac{ESS_R - ESS_U}{df_R - df_U}}{\dfrac{ESS_U}{df_U}}$$

where df stands for degrees of freedom. f measures the additional explanatory power of the unrestricted compared to the restricted equations. With true restrictions, f will equal zero since the constrained equation contains the true parameter values. In such a case, calculating estimates of these true values cannot lead to greater explanatory power for the unconstrained regression. In any sample, when the calculated statistic lies below the critical value, the restrictions hold.

The test for aggregation bias has the null hypothesis that each coefficient parameter takes an identical magnitude for every firm. Formally, $\beta_{11} = \beta_{12} = \ldots = \beta_{1i}$ for all h which implies that the 95 firms behave homogeneously with respect to the regression coefficients and, thus, no aggregation bias exists. This result follows since under the null hypothesis $\frac{1}{N} \sum_i \beta_{hi} = \beta_{hi} = \beta_h$ where β_h is the macroparameter from the incorrect macrorelation. In this case, the correct and incorrect macrorelations have identical parameters.

The test statistic must account for the correlation between the disturbances of individual firm regressions. To accomplish this result, the two-stage Zellner (2SZ) technique is applied to each of the five specifications in both unrestricted and restricted forms. The unrestricted form assumes the alternative hypothesis and, for each specification, requires the estimation of the system of 95 individual firm microrelations. The restricted form imposes the information contained in the null hypothesis on the same system of 95 microrelations before estimation occurs. The resulting f statistic measures the additional explanatory power when adding the restrictions. A sufficiently large value for f would lead to rejection of the null hypothesis and imply the existence of bias. According to Zellner (1962), the distribution of f approximates that of $f_{q,n-m}$ where q is the number of restrictions, n stands for the number of observations, and m represents the number of independent variables. Implementation of both tests and construction of aggregation bias will occur in Chapter IV.

Summary. Samples containing firm level data allow a logical sequence in the testing of money demand models. The first step tests the validity of using a macrorelation to evaluate micro economic behavior. The absence of aggregation bias implies pooling firm level data to improve the estimators. However, individual firm regressions must be used when large amounts of bias exist. After establishing the plane on which estimation will occur, two samples exist to evaluate models of money demand.

Summary

Chapter III has applied an econometric analysis to firms' money demand. This methodology first develops testable specifications, then selects econometric techniques, and finally gathers a set of facts.

Eight inventory, one portfolio, and three production specifications evolved as follows. Existing functions from previous literature were adjusted by the discussion of money's price to ensure useful model comparisons. Variables were chosen to accurately measure relevant theoretical concepts contained in the adjusted equations. Combining these steps resulted in the 12 testable specifications.

Six advanced econometric techniques are necessary to form statistics measuring these specification's performance. The Box-Cox analysis selects specific forms. Ridge regression diagnoses and corrects for multicollinearity. Two-stage Zellner and three-stage least squares estimate systems of equations. Finally, the covariance and error components model adjust for omitted variables and/or pooled data.

Two samples provide arenas for the models' performances. Both the quarterly and annual data bases contain firm level observations on all variables in every specification. The latter characteristic provides for accurate comparative performances. The former property allows tests for aggregation bias to determine the correct plane for generating performance statistics.

Combining the three pieces of the methodology gives a systematic approach to evaluating firms' money demand. Each specification embodies the unique aspects of its corresponding theory. The performance of these equations exists in statistics generated by combining the samples with the econometric techniques. Chapter IV carries out that task.

CHAPTER IV

ECONOMETRIC RESULTS

Introduction

Chapter IV evaluates the three money demand models by implementing the econometric analysis developed in Chapter III. This examination occurs in the following sequence.

The quarterly sample implies a set of specifications for the three theories. After listing these equations, different econometric methods combine with the data until the superior technique is discovered. The statistics generated by this technique provide the basis for evaluating the models. Both estimation and prediciton criteria exist to make the examination. Then the same sequence repeats for the annual sample. Finally, a summary section presents the results and implications from both samples.

Quarterly Sample

Specifications

The quarterly sample allows for the estimation of and predictions from the following specifications. Table I contains the definitions of all symbols.

Traditional Inventory Model (TIM):

$$LN(M) = \beta_0 + \beta_1\ LN(S) + \beta_2\ LN(FFR) + \varepsilon_1 \qquad (4\text{-}1)$$

TABLE I

VARIABLES[1]

Symbol	Definition
CA	Real current assets[2]
CBR	Ten-year corporate bond rate
CR	Real cash receipts
EQR	Equity returns for individual firms
FFR	Federal funds rate
M	Real money balances
PPC	Product price index for specific commodities
PPI	Percentage change in the product price index
Q	Physical output
RPV	Inverse of the volume of repurchase agreement trading
RR	Price of capital
S	Real sales
TBR	Treasury bill rate
V	Real cash flow variance
WPI	Wholesale price index
WR	Wage rate for specific commodities

[1]See Appendix D for a detailed description.

[2]Real values are nominal divided by the wholesale price index unless stated otherwise.

Deterministic Inventory Model (DIM):

$$LN(M) = \beta_0 + \beta_1 LN(CR) + \beta_2 LN(FFR) + \varepsilon_2 \tag{4-2}$$

Portfolio Model (PM):

$$LN(M) = \beta_0 + \beta_1 LN(CA) + \beta_2 LN(FFR) + \beta_3 LN(CBR) + \beta_4 LN(EQR) + \beta_5 LN(PPI) + \varepsilon_3 \tag{4-3}$$

Transactions Cost Model (TCM):

$$LN(M) = \beta_0 + \beta_1 LN(PPC) + \beta_2 LN(FFR) + \beta_3 LN(EQR) + \beta_4 LN(WR) + \beta_5 LN(RR) + \varepsilon_4 \tag{4-4}$$

Neoclassical Model (NM):

$$LN(M) = \beta_0 + \beta_1 LN(Q) + \beta_2 LN(FFR) + \beta_3 LN(WR) + \beta_4 LN(RR) + \varepsilon_5 \tag{4-5}$$

Stochastic Inventory Model (SIM):

$$LN(M) = \beta_0 + \beta_1 LN(V) + \beta_2 LN(FFR) + \varepsilon_6 \tag{4-6}$$

Stock Adjustment Inventory Model (SAIM):

$$LN(M) = \beta_0 + \beta_1 LN(CR) + \beta_2 LN(FFR) + \beta_3 LN(M_{-1}) + \varepsilon_7 \tag{4-7}$$

The log liner forms of Equations (4-1), (4-2), (4-6), and (4-7) stem directly from theory. The Box-Cox analysis used pooled data to select the specifications for Equations (4-3), (4-4), and (4-5). Table II contains the test results. Models (4-6) and (4-7) failed to provide adequate explanations of money demand as indicated by initial estimation. The stochastic inventory model had an adjusted R-squared of only .20. The standard f test for joint inclusion of the regressors

TABLE II

ORDINARY LEAST SQUARES (QUARTERLY)

Equation	Coefficient Estimate[1]	S	Box-Cox Test	Adjusted Durbin-Watson Test	Bartlett's Test
Equation (4-1) - TIM[2]		0.90826		0.42953	8.43215*
INT[3]	-2.1824				
	(0.1982)				
LN(S)	1.1275				
	(0.0224)				
LN(FFR)	-0.2923				
	(0.0796)				
Equation (4-2) - DIM		0.87819		0.45137	9.09241*
INT	-1.8802				
	(0.1873)				
LN(CR)	1.0526				
	(0.0199)				
LN(FFR)	-0.6098				
	(0.0774)				
Equation (4-3) - PM		0.83564	2.64341*	0.42639	4.74573*
INT	-1.2280				
	(0.8351)				
LN(CA)	1.0798				
	(0.0192)				
LN(FFR)	-0.1554				
	(0.2484)				
LN(CBR)	-0.3141				
	(0.6108)				
LN(EQR)	-0.4895				
	(0.3475)				
LN(WPI)	-0.0510				
	(0.1049)				
Equation (4-4) - TCM		1.48750	3.48416*	0.49708	10.4858*
INT	-4.0874				
	(0.9877)				
LN(PPC)	1.1295				
	(0.1725)				
LN(FFR)	-0.0186				
	(0.2128)				
LN(EQR)	-0.6974				
	(0.4228)				
LN(WR)	0.2616				
	(0.0702)				
LN(RR)	2.6469				
	(0.7195)				
Equation (4-5) - NM		0.87018	1.99352*	0.46277	9.09714*
INT	-3.6385				
	(0.2599)				
LN(Q)	1.0471				
	(0.0231)				
LN(FFR)	-0.1889				
	(0.0781)				
LN(WR)	1.2662				
	(0.1309)				
LN(RR)	0.1443				
	(0.0439)				

[1]Standard error in parenthesis: S symbolizes standard error of regression. Asterisk (*) designates acceptance of the null hypothesis at 1% significance level. For example, no asterisk on a coefficient estimate indicates significance. One percent (1%) was selected to indicate the strength of the conclusions. The test results hold at this 5% level.

[2]All abbreviations correspond to those in Table I.

[3]Intercept term.

rejected the hypothesis that their influence was significant. Finally, all independent variables had insignificant coefficient estimates as given by standard t tests. Estimation of the stock-adjustment inventory model rendered an insignificant coefficient for the lag of money balances. This result implies that the stock adjustment inventory model reduces to Specification (4-2). Because of these findings, Equations (4-6) and (4-7) are not included in the empirical analysis that follows.

Estimation

Ordinary Least Squares. Given the five specifications, the next task is to find unbiased and efficient estimators. The ordinary least squares estimates for each equation using pooled observations appear together in Table II along with various test results. These estimates must undergo examination for autocorrelation and heteroscedasticity. With respect to the former, the adjusted Durbin-Watson test (see Appendix E for test procedure) indicates the presence of first order autocorrelation in every equation. The existence of heteroscedasticity can be determined by Bartlett's test (see Appendix G for test procedure). For each specification the procedure cannot reject a homoscedastic variance. Thus, OLS does not require any adjustment for this condition. However, an efficiency improvement in the OLS estimates will occur by correcting for the autocorrelation.

Generalized Difference Model. Using pooled data necessitates the following adjustment procedure for every equation. Ordinary least squares is applied to each of the 95 firms individually. For every regression, the Cochrane-Orcutt method employs the residuals to calculate an estimate of the first order autocorrelation coefficient ($\hat{\rho}_i$).

Each of the 95 $\hat{\rho}_i$ transforms the 30 observations for its corresponding firm by generalized differencing. Once transformed, all observations are pooled and subjected to OLS. The entire procedure must be repeated for each equation individually, since the $\hat{\rho}_i$ for the i^{th} firm will differ for different specifications. In total, 95 times five $\hat{\rho}_i$ are calculated and each of the five sets transforms data for its corresponding equation. This process, called the generalized difference model (GDM), generates the estimates and tests appearing in Table III. These findings must undergo the adjusted Durbin-Watson and Bartlett's tests as a check on disturbance term patterns. In all cases, the examinations show that the disturbances conform to the classical linear regression model. Thus, the GDM estimates provide a basis for further testing.

The next step re-examines the specifications of Equations (4-3), (4-4), and (4-5). Such redundancy was considered necessary since the Box-Cox analysis uses the data (now transformed) to find an optimal specification. Table III contains the test statistics (for the GDM when pooling all available observations) which imply log-linear forms for Equations (4-3), (4-4), and (4-5). Thus, when pooling observations, all five equations take linear in the natural logarithm specifications. Yet, pooling should only occur when the models contain a statistically insignificant amount of aggregation bias. Having firm level data allows the rare possibility of computing and testing the amount of bias in each specification.

The results of the computations for each equation appear in Table III. The bias from both corresponding and noncorresponding microparameters is small for every variable in each specification. The range of

TABLE III

GENERALIZED DIFFERENCE MODEL (QUARTERLY)

	Estimation					Aggregation							Multicollinearity			
													k=.20		k=.25	
Equation	Coefficient Estimate[1]	S	Box-Cox Test	Adjusted Durbin-Watson Test	Bartlett's Test	Macro-Parameter Estimate	Total Aggregation Bias[3]	Corresponding Bias	Non-corresponding Bias	True Parameter Value	Deviation From Expectation[4]	F-Test	Coefficient Estimate	S	Coefficient Estimate	S
Equation (4-1) - FIM[2]		0.52534		2.14107*	5.68128*							1.3488*				
INT[5]	-1.8933 (0.1886)															
LN(S)	1.0477 (0.0243)					0.93820	0.01644	0.02631	-0.00987	0.92456	-0.00180					
LN(FFR)	-0.2298 (0.0590)					-0.26854	0.00478	-0.00678	0.01156	-0.26751	-0.00375					
Equation (4-2) - DIM		0.51954		2.13461*	4.79612*							1.2975*				
INT	-1.3235 (0.2081)															
LN(CR)	0.9139 (0.0304)					0.87915	0.01419	0.02644	-0.01225	0.87015	-0.00509					
LN(FFR)	-0.5355 (0.0603)					-0.48026	0.01299	-0.01013	0.02312	-0.49711	0.00386					
Equation (4-3) - PM		0.48228	2.94802*	2.19965*	4.51386*							1.1063*				
INT	-0.7555 (0.5750)															
LN(CA)	1.0434 (0.0303)					0.99026	0.01252	0.02817	-0.01565	0.98517	-0.00743					
LN(FFR)	-0.1523 (0.1389)					-0.19731	-0.01243	-0.00341	-0.00902	-0.21457	0.00483		-0.1448	0.49381	-0.1368	0.50052
LN(CBR)	-0.7879 (0.3303)					-0.61833	-0.03909	-0.01123	-0.02786	-0.66548	0.00806		-0.7626		-0.7484	
LN(EQR)	-0.2638 (0.2833)					-0.31064	-0.01839	-0.00591	-0.01248	-0.32316	-0.00587		-0.2453		-0.2237	
LN(PPI)	0.0232 (0.0549)					0.02951	-0.00064	0.00077	-0.00141	0.02781	0.00106		0.0204		0.0195	
Equation (4-4) - TCM		1.03155	2.07414*	2.00315*	7.19128*							1.4866*				
INT	0.8171 (0.9153)															
LN(PPC)	-0.1248 (0.0141)					-0.15006	-0.00038	-0.00300	0.00262	-0.14761	-0.00283					
LN(FFR)	-0.3145 (0.2817)					-0.36811	0.00499	-0.00811	0.01310	-0.36731	0.00419		-0.3077	1.12651	-0.2985	1.15718
LN(EQR)	0.4538 (0.3901)					0.45219	-0.00797	0.00761	-0.01558	0.43128	0.01294		0.4338		0.4066	
LN(WR)	1.8761 (1.1030)					1.60812	0.07132	0.03113	0.04019	1.69256	-0.01312		1.6891		1.6479	
LN(RR)	2.3458 (1.5132)					2.00719	0.11381	0.05561	0.05820	2.10369	0.01731		2.1411		2.0038	
Equation (4-5) - NM		0.51693	0.00000*	2.13816*	6.02185*							1.3221*				
INT	-3.0050 (0.2986)															
LN(Q)	0.9276 (0.0320)					0.95797	0.03989	0.02505	0.01484	1.00362	-0.00576					
LN(FFR)	-0.1296 (0.0636)					-0.17955	0.00367	-0.00287	0.00654	-0.18123	0.00536		-0.1134	0.52033	-0.1005	0.52281
LN(WR)	0.9951 (0.1755)					0.91176	0.05051	0.02238	0.02813	0.95661	0.00666		0.9467		0.9255	
LN(RR)	0.1969 (0.0544)					0.20374	0.01125	0.00301	0.00824	0.22810	-0.01311		0.1873		0.1664	

[1]See note 1, Table II.

[2]See note 2, Table II.

[3]Sum of corresponding and non-corresponding bias.

[4]Equals the macroparameter estimate minus the sum of total bias and the true parameter value.

[5]Intercept term.

total bias stretches from three to eight percent of the respective parameter estimate while its average amount for all variables is less than five percent. These findings imply a small loss of accuracy when aggregating. The next step examines just how small the magnitude of bias is with a test of significance.

The procedure results in the f statistic values appearing in Table III. For every specification the test renders support for the absence of aggregation bias, which justifies pooling the data. Using all observations, the generalized difference model can generate unbiased and efficient estimates. Yet, more accurate estimates may exist if a model exhibits multicollinearity.

The next task examines Specifications (4-3), (4-4), and (4-5) for such correlation. The portfolio model may exhibit multicollinearity between the various rates of return while input prices could be sources in either the transactions cost model or neoclassical model. Applying ridge regression to the transformed data with k values of .10 and .25 gives the results in Table III. None of the suspected variables' estimates deflate substantially from their OLS values at either k value. Such results imply the absence of multicollinearity for all three specifications. The GDM estimates provide precise and stable measures of each variable's influence on money demand. Yet, pooling the data suggests one final approach to improving these estimates.

Pooling Adjusted Methods. Gains may occur when adjusting for any differences in the intercept terms of individual cross-sections and time series regressions. Two econometric techniques exist that can accomplish this task.

Table IV contains the results from the Error Components Model (ECM)

TABLE IV

POOLING ADJUSTED MODELS (QUARTERLY)

	Covariance Model						Error Components Model	
Equation	Coefficient Estimate	S	Adjusted Durbin-Watson Test	Bartlett's Test	Zero Correlation Coefficient Test	Joint Inclusion f Test	Coefficient Estimate	S
Equation (4-1) - TIM[2]		0.50062	2.13503*	4.84101*	0.1154*	184.231		
INT[3]	2.7744				(0.1618)		0.0085	0.51679
	(0.5603)						(0.3849)	
LN(S)	0.4483						0.7197	
	(0.0911)						(0.0656)	
LN(FFR)	-0.2283						-0.2170	
	(0.0606)						(0.0496)	
Equation (4-2) - DIM		0.49967	2.13819*	3.98610*	0.1004*	185.606		
INT	2.3308				(0.1369)		-0.2892	0.51073
	(0.5931)						(0.3518)	
LN(CR)	0.4299						0.7597	
	(0.0989)						(0.0582)	
LN(FFR)	-0.4626						-0.4581	
	(0.0261)						(0.0561)	
Equation (4-3) - PM		0.46559	2.23040*	3.71844*		205.711		
INT	-4.0389						-1.5412	0.46796
	(0.8374)						(0.7542)	
LN(CA)	1.5011						1.2137	
	(0.1188)						(0.0613)	
LN(FFR)	-0.3314						-0.0334	
	(0.1413)						(0.1749)	
LN(CBR)	-0.4811						-0.5772	
	(0.3428)						(0.5050)	
LN(EQR)	-0.7343						-0.5971	
	(0.2035)						(0.3302)	
LN(PPI)	-0.0588						-0.0470	
	(0.0596)						(0.0997)	
Equation (4-4) - TCM		0.99883	2.14542*	5.86177*	0.3889	100.141		
INT	-3.4221				(0.1366)		-3.0166	1.00682
	(0.8531)						(0.8477)	
LN(PPC)	0.1826						0.8861	
	(0.2159)						(0.2168)	
LN(FFR)	0.1088						-0.1172	
	(0.0951)						(0.2014)	
LN(EQR)	0.7110						-0.5129	
	(0.2043)						(0.3872)	
LN(WR)	-0.1639						-0.1445	
	(0.3223)						(0.0688)	
LN(RR)	1.4852						1.6438	
	(1.3321)						(1.3127)	
Equation (4-5) - NM		0.49819	2.13132*	3.98612*	0.1248*	180.628		
INT	0.0479				(0.1932)		-2.1779	0.51357
	(1.2552)						(0.6278)	
LN(Q)	0.3095						0.6365	
	(0.1041)						(0.0671)	
LN(FFR)	-0.3081						-0.0672	
	(0.1172)						(0.0688)	
LN(WR)	1.9594						1.6695	
	(0.5035)						(0.3541)	
LN(RR)	0.4795						0.2558	
	(0.3430)						(0.1449)	

[1]See note 1, Table II.

[2]See note 2, Table II.

[3]Intercept term.

using all raw data observations. Two findings stand out. First, the standard error of regression (S) for each model is reduced compared to the OLS model. The improved fit implies that cross-sectional and time series differences do exist and, thus, the ECM estimates have improved accuracy compared to those from OLS. Second, most coefficient estimates are smaller than their OLS counterparts. Allowing for differences among firms and across time highlights an overprediction made by OLS which could hide the presence of economies of scale. While these estimates give improvements over OLS, they assume that the time series correlation between this period's disturbance and that of a previous period remains fixed no matter how far in the past the previous period lies. However, the raw data exhibits first order autocorrelation which postulates that this period's disturbance has a weaker correlation with disturbances which occurred further in the past. For this reason, the ECM will not give the most efficient estimators possible. Fortunately, the second technique can incorporate the correct disturbance pattern.

The covariance model (COVM) estimates and corresponding tests appear in Table IV and have two major implications. First, the f test for joint significance of all dummy variables shows that the COVM represents a significant improvement over the generalized difference model (GDM). Second, this conclusion is reinforced by the lower S values for the COVM compared to the GDM. The COVM estimates represent the most efficient, unbiased ones available. Consequently, they provide the preferred basis for comparing the specifications.

Evaluation Based on Estimation

The following criteria will be used to discriminate between the estimated models: magnitude of S^2 (estimated variance of the specification), joint significance test of all regressors, individual significance tests for each regressor, and signs and magnitudes of coefficient estimates.

The broadest indicator of a specification's performance measures its explanatory power. The magnitude of S^2 shows the amount of variation in the dependent variable not accounted for by all regressors. Theil (1971) has shown that when a set of specifications explains the same dependent variable, the correct specification cannot have a larger expected value of S^2 than an incorrect one.[1]

The values for S^2 from GDM on each specification appear in Table IV. Inspection renders the following ranking from smallest (best) to largest (worse): portfolio model, neoclassical model, deterministic inventory model, traditional inventory model, and transactions cost model. In order to deem Equation (4-3) superior, its S^2 must have a significantly lower value than the alternatives. Establishing this property cannot rely on the f test for equal variances since it assumes independence between the S^2 generated by different equations. The correct technique is the test for a zero correlation coefficient, outlined by Granger and Newbold (1973).

The computations for the correlation coefficients between each model and the portfolio model and their standard errors appear in Table IV. Only the transactions cost model has inferior explanatory power.

---

[1]For the five models under study, generalized differencing ensures that each specification explains a slightly different dependent variable. This result occurs since the estimated autocorrelation coefficient for any firm differs for each specification. Yet, S^2 of the GDM equations still provides accurate comparisons as proved by Goldberger (1965).

No significant difference exists in the performance of the other four specifications at the 1% significance level.

Although relative performance is an important measure, interest also lies in the absolute ability of these models to explain money demand. This ability can be examined using the f test for joint inclusion of the regressors. Inspection of Table IV shows that all specifications contain regressors which taken together have a significant influence on money demand.

The next step presents tests of the explanatory power of individual regressors using the t test for coefficient significance. Looking at the results in Table IV shows that only the two inventory specifications have significant estimates for all regressors. The portfolio model contains two insignificant estimates; those for the corporate bond rate and the per-centage change in the product price index.[2] In the transactions cost model, the product price index for specific commodities, the wage rate, and the price of capital have insignificant coefficients. Finally, the neoclassical model shows insignificant explanatory power for the price of capital. These results lift the two inventory models above the alternatives.

The final criteria of the estimate's signs and magnitudes make further distinctions. The traditional inventory model has estimates of the correct sign for both regressors. t tests show that while the sales elasticity does not differ from one-half, the interest rate coefficient differs from a negative one-half. However, the deterministic inventory model has cash receipt and interest rate elasticity estimates not different from one-half and negative one-half, respectively. These

---

[2]The insignificance of PPI is not a strike against the portfolio model. It indicates that the equation is homogenous of degree one.

findings support cash receipts as the proper transactions variable in an inventory model. The significant regressor estimates in the portfolio model take the correct signs. The elasticity of current assets exceeds one in denial of economies of scale. The federal funds rate and equity returns have an inverse affect on money balances. The transactions cost model's only two significant coefficients have incorrect signs. In contrast to expectations, changes in either the federal funds rate or equity returns change money demand in the same direction. Finally, all significant estimates for the neoclassical model take the correct signs. The elasticity of real output lies below one implying economies of scale. The negative coefficient for the federal funds rate means an inverse relationship with cash holdings. The positive estimate for the wage rate indicates that money balances substitute for labor in the production process.

In summary, the only model supported by all estimation criteria is the deterministic inventory model. Also, its explanatory power does not differ significantly from the alternatives. These findings raise Equation (4-2) above the other models as an explanation of cash holdings. Yet, in certain circumstances, a model that explains well may predict poorly outside of the sample used to construct estimates. Since policy decisions depend on the forecasting ability of these models, prediction performance is an important independent means to evaluate the models.

Prediction

Formation of a Forecast Series. Predictions from the five specifications use coefficient estimates from the COVM since they are minimum variance and unbiased. These five models make forecasts after an adjustment for serial correlation,[3] of one observation per firm on the dependent variable for each quarter from 1979-2 through 1980-1. This prediction sample permits each specification to make 380 one step ahead forecasts. The resulting set of predictions, called a forecast series, can be evaluated by both absolute and relative criteria.

Evaluation of a Forecast Series. The absolute prediction criteria include mean square error (MSE) and simple regressions. Relative criteria come from composite predictors.

Absolute Criteria. The fundamental criteria of MSE measures the inaccuracy of the forecast series. Computationally, it equals

$$\text{MSE} = \frac{1}{T}\sum_{t}^{T} e_t^2$$

where e represents the difference between actual and predicted magnitudes, and t sums over all forecasts. MSE gives the prediction counterpart to S^2 formed for estimation. In fact, the two are computed identically except that S^2 uses within sample estimates of the dependent variable while MSE employs dependent variable values outside the sample.

Table V contains the results of calculating MSE for every specification. Inspection indicates the following ranking from most accurate to least accurate predictor: the portfolio model, the neoclassical model, the deterministic inventory model, the traditional inventory model, and the transactions cost model. To assert the superiority of the portfolio model requires that its MSE lies significantly below the

---

[3]This adjustment insures that all equations predict the same dependent variable.

MSE for competing models. The test for zero correlation coefficient can make this determination (see Appendix G for test procedure).

The results of computations for r_{13}, r_{23}, r_{43}, and r_{53} and their standard errors appear in Table V. Based on those values, only the transactions cost model gives inferior forecasts. No difference exists in the accuracy of the other four specifications.

The error series not only provides this measure of a specification's forecast accuracy, tests of its mean and autocorrelation can reveal underlying properties of the predictor. If the error series has zero mean, the corresponding predictor makes unbiased forecasts. Lack of autocorrelation implies that no useful information known at the time of a forecast was wasted. If such time correlation exists, it could revise the forecast improving its accuracy. Any unbiased predictor with a non-autocorrelated error series is called an optimal predictor.

Two tests can establish whether any of the five specifications generate optimal predictors. A t test can examine the error series for zero mean. The t statistic divides the sample estimate of the mean by its standard error corrected for the number of observations. If the calculated statistic exceeds its critical value, the mean differs from zero. Table V contains the results of these calculations for each model. Implementing the tests indicates that all five specifications generate unbiased forecasts of the dependent variable. The second test explores the error series' time correlation properties. An adjusted Von-Neuman ratio test makes this determination. (See Appendix G for test procedure.) The results of the test for all specifications appear in Table V and, in each case imply the absense of autocorrelation. Combining this finding with the property of unbiasedness shows that all

TABLE V

PREDICTION RESULTS (QUARTERLY)

Model	Error Series				Predictor Series: Simple Regressions				Composite Predictors Unconstrained Then Constrained[3]				
	MSE	Test for Zero Correlation Coefficient[1]	Test for Zero Mean	Non-Nueman Ratio	INT	P_i	S	Test for Zero Correlation Coefficient	P_1 P_2 P_3 P_4 and P_5	P_3 and P_1	P_3 and P_2	P_3 and P_4	P_3 and P_5
Model (4-1) TIM[2]	0.65130	0.1001* (0.1949)	0.1284* (0.3601)	1.88561*	0.2145 (0.2248)	0.8964 (0.1073)	0.63481	0.0899* (0.1138)	0.2451* (0.3009)	0.1243* (0.1481)			
									0.2811* (0.2776)	0.1023* (0.1014)			
Model (4-2) DIM	0.64247	0.0894* (0.1092)	0.1060* (0.3465)	1.89528*	0.1882 (0.2091)	0.9151 (0.0943)	0.62884	0.0754* (0.0843)	0.1879* (0.2968)		0.1051* (0.1642)		
									0.1746* (0.2441)		0.0986* (0.1176)		
Model (4-3) PM	0.60640	---	0.0858* (0.2760)	1.95582*	0.0645 (0.1891)	0.9644 (0.0900)	0.59325	---	0.1561* (0.2201)	0.8823 (0.1264)	0.8916 (0.1179)	0.8657 (0.0858)	0.9018 (0.1066)
									0.1601* (0.2109)	0.8977 (0.1014)	0.9014 (0.1176)	0.8351 (0.2418)	0.9210 (0.1093)
Model (4-4) TCM	1.27563	0.3924 (0.1363)	0.2882* (0.5865)	1.87113*	0.5478 (0.3142)	0.7314 (0.2063)	0.98792	0.4234 (0.1485)	0.2811* (0.3792)			0.2010* (0.1905)	
									0.2955* (0.2109)			0.1649* (0.2418)	
Model (4-5) NM	0.63925	0.0806* (0.1074)	0.0983* (0.2835)	1.90553*	0.1653 (0.2001)	0.9239 (0.9020)	0.61004	0.0652* (0.0944)	0.0913* (0.2811)				0.0897* (0.9018)
									0.0887* (0.3199)				0.0790* (0.1093)
S[4]									0.50561	0.56381	0.55812	0.56177	0.55743
									0.52009	0.57443	0.56664	0.57842	0.56383

[1]Standard errors in parenthesis: Asterisk (*) designates acceptance of the null hypothesis at 1% significance level.

[2]See note 2, Table II.

[3]Coefficient estimates for the listed predictors read down the column corresponding to their model number. The upper estimates are for unrestricted form while the lower estimates are those from constrained forms.

[4]Standard error for composite predictor in the same column; upper number for unconstrained form, lower number for constrained form.

five specifications generate optimal forecast series. To solidify this finding, an alternative way to examine for optimality uses the predictor series not its corresponding error series.

Cooper and Nelson (1975) outline the technique which involves estimating a simple regression of the predicted values (P) on actual outcomes (M). The exact form of each equation is

$$M_t = \alpha_0 + \alpha_1 P_t + u_t$$

where u represents a classical linear regression model disturbance term. As shown by Mincer and Zarnowitz (1969), an optimal forecaster will have estimates of α_0 and α_1 not significantly different from zero and one, respectively. t tests can validate these conditions.

For the five equations, the relevant t statistics appear in Table V. All specifications have estimates of α_0 and α_1 not significantly different from zero and one, respectively. This procedure lends further support to the contention that every model makes optimal forecasts. Since no distinction between them exists on the basis of optimality, the next step examines a measure of accuracy from these simple regressions.

As Granger and Newbold (1973) have shown, the most accurate predictor, out of some set of optimal predictors, has the smallest S^2 from the simple regressions. Looking at these values ranks the specifications as follows: the portfolio model, the neoclassical model, the deterministic inventory model, the traditional inventory model, and the transactions cost model. Concluding the superiority of the portfolio model necessitates that its S^2 is significantly less than that of alternative models. The zero correlation coefficient test can make this determination.

The values for r_{13}, r_{23}, r_{43}, and r_{53} appear, with their standard errors, in Table V. The transactions cost model gives inferior performance compared to the alternative models. No difference exists in the accuracy of the other four specifications.

In summary, using either the error series or the predictor series leads to identical findings. All specifications make optimal predictions. The accuracy of the portfolio, traditional inventory, deterministic inventory, and neoclassical models are statistically identical to each other and significantly better than that of the transactions cost model. Thus, no distinction between the Equations (4-1), (4-2), (4-3), and (4-5) comes from absolute prediction measures. Any difference must occur based on criteria from relative forecast performance within a composite predictor.

Relative Criteria. A composite predictor is a multivariate regression on actual outcomes of the dependent variable using at least two prediction series as explanatory variables. In such a regression, the coefficient estimates measure the partial derivative of the dependent variable with respect to the regressor. If the estimate differs from zero, the corresponding regressor explains a significant amount of the variation in the dependent variable, given that the influence of other regressors has been accounted for. In other words, a significant coefficient implies that a predictor series adds useful information not contained in the other forecasts. An insignificant estimate would mean the predictor has no relative explanatory power.

The composite predictor containing all specifications appears as,

$$M = \alpha_1 P_1 + \alpha_2 P_2 + \alpha_3 P_3 + \alpha_4 P_4 + \alpha_5 P_5 + \varepsilon \qquad (4\text{-}8)$$

where M represents actual outcomes of money balances, P_i stands for the prediction series from the i^{th} specification, and ε is a classical linear regression model disturbance. When all forecasts make unbiased predictions the regressor coefficients must sum to unity.[4]

The results for both unconstrained and constrained equations appear in Table IV. Since the f statistic does not exceed its critical value, the five coefficients sum to unity. This result reconfirms the lack of bias in the predictions made by the five specifications. Also, all coefficient estimates in both equations have t statistics less than the critical value. No single specification makes a marginally significant contribution to the remaining group. Yet, some doubt arises about the validity of these t tests. Because the five predictor series move together, multicollinearity exists in the $X'X$ matrix used to compute these estimates. If this condition inflates standard errors, the calculated t statistics may appear too small. An alternative set of composite predictors was formed to avoid this potential problem.

Four regressions were estimated as,

$$M = \alpha_3 P_3 + \alpha_i P_i + \varepsilon \tag{4-9}$$

---

[4]Testing for this property requires estimating an unconstrained form, as Equation (4-9), then a constrained form which assumes $\alpha_3 = 1 - \alpha_1 - \alpha_2 - \alpha_4 - \alpha_5$. The constrained equation results from substituting this restriction into Equation (4-9) to obtain,

$$e_3 = \alpha_1(P_1 - P_3) + \alpha_2(P_2 - P_3) + \alpha_4(P_4 - P_3) + \alpha_5(P_5 - P_3) + \varepsilon$$

where e_3 equals the error series from the portfolio specification. The test procedure uses the error sum squares of the two regressions in forming an f statistic. Its magnitude measures the additional explanatory power when removing the restrictions. Thus, a small f, less than the critical value, validates the restrictions.

where i = 1, 2, 4, 5. To test unbiasedness, a constrained form ($\alpha_3 = 1 - \alpha_i$) appearing as,

$$e_3 = \alpha_i(P_i - P_3) + \varepsilon \tag{4-10}$$

also underwent estimation. If P_3 contains all the useful information of P_i, then, in Equation (4-9), α_i will equal zero and α_3 unity. When P_i has useful information independent of P_3, its coefficient in Equation (4-9) will differ from zero implying that α_3 differs from one. With respect to Equation (4-10), α_i will equal zero if P_i adds explanatory power beyond that of P_3 but will differ from zero when P_i gives independent information.

The results for all eight regressions appear in Table V. For the unconstrained form, no instance occurred when α_i and α_3 differed from zero and unity as established by t tests. This result implies that P_3 subsumes the alternative predictors. In constrained form, t tests show that none of the α_i differed from zero which confirms the prior finding. Finally, none of the calculated f values exceeded the critical magnitude, indicating that the predictors make unbiased forecasts.

In summary, relative prediction criteria make no distinctions between the five models. None make a marginally significant contribution to either the group or to predictor three. Absolute prediction criteria show the following. All specifications make optimal forecasts. The transactions cost model gives less accurate predictions than the alternatives. Finally, the remaining four specifications give predictions of equal accuracy.

Summary

Based on both prediction and estimation criteria, only the deterministic inventory model finds total support. Only this specification passes all the examinations made from estimation. Since its predictions do not differ significantly from the alternatives, the deterministic inventory model represents the superior specification.

Annual Sample

Overview

In order to substantiate this conclusion, the annual sample was employed in making estimates and predictions. This data base also allowed the examination of more complex models and the cash management technique literature. The exposition of the findings based on annual data follows the same pattern as with the quarterly sample. Also, many of the same test procedures were used. Thus, details of the exposition and tests will not be repeated in the following analysis.

Specifications

The annual sample gives a basis for testing the following specifications. The symbols are defined in Table I.

Traditional Inventory Model (TIM):

$$LN(M) = \beta_0 + \beta_1 \, LN(S) + \beta_2 \, LN(FFR) + \varepsilon_1 \tag{4-11}$$

Deterministic Inventory Model (DIM):

$$LN(M) = \beta_0 + \beta_1 \, LN(CR) + \beta_2 \, LN(FFR) + \varepsilon_2 \tag{4-12}$$

Portfolio Model (PM):

$$\begin{aligned} LN(M) = \beta_0 &+ \beta_1 \, LN(CA) + \beta_2 \, LN(FFR) + \beta_3 \, LN(CBR) \\ &+ \beta_4 \, LN(EQR) + \beta_5 \, LN(PPI) + \varepsilon_3 \end{aligned} \tag{4-13}$$

Transactions Cost Model (TCM):

$$LN(M) = \beta_0 + \beta_1\ LN(PPC) + \beta_2\ LN(FFR) + \beta_3\ LN(EQR) + \beta_4\ LN(WR) + \beta_5\ LN(RR) + \varepsilon_4 \quad (4\text{-}14)$$

Neoclassical Model (NM):

$$LN(M) = \beta_0 + \beta_1\ LN(Q) + \beta_2\ LN(FFR) + \beta_3\ LN(WR) + \beta_4\ LN(RR) + \varepsilon_5 \quad (4\text{-}15)$$

Translog Cost Model (TCM):

$$LN(M) = \beta_0 + \beta_1\ LN(Q) + \overline{\beta}_2\ LN(FFR) + \overline{\beta}_3\ LN(WR) + \overline{\beta}_4\ LN(RR) + \varepsilon_6 \quad (4\text{-}16)$$

Complete Traditional Inventory Model (CTIM):

$$LN(M) = \beta_0\ LN(RPV) + \beta_1\ LN(S) + \beta_2\ LN(FFR) + \varepsilon_7 \quad (4\text{-}17)$$

Complete Deterministic Inventory Model (CDIM):

$$LN(M) = \beta_0\ LN(RPV) + \beta_1\ LN(CR) + \beta_2\ LN(FFR) + \varepsilon_8 \quad (4\text{-}18)$$

Complete Stochastic Inventory Model (CSIM)

$$LN(M) = \beta_0\ LN(RPV) + \beta_1\ LN(V) + \beta_2\ LN(FFR) + \varepsilon_9 \quad (4\text{-}19)$$

Complete Stock Adjustment Inventory Model (CSAIM):

$$LN(M) = \beta_0\ LN(RPV) + \beta_1\ LN(CR) + \beta_2 LN(FFR) + \beta_3\ LN(M_{-1}) + \varepsilon_{10} \quad (4\text{-}20)$$

Cash Management Technique Inventory Model (CMTIM):

$$LN(M) = \beta_0\ LN(RPV) + \beta_1\ LN(CR\ ) + \beta_2\ LN(FFR) + D_{66}\beta_0\ LN(RPV) + D_{66}\beta_1\ LN(CR) + D_{66}\beta_2\ LN(FFR) + D_{69}\beta_0\ LN(RPV) \quad (4\text{-}21)$$

$$+ D_{69}\beta_1 \text{ LN(CR)} + D_{69}\beta_2 \text{ LN(FFR)} + D_{73}\beta_0 \text{ LN(RPV)}$$

$$+ D_{73}\beta_1 \text{ LN(CR)} + D_{73}\beta_2 \text{ LN(FFR)} + \varepsilon_{11}$$

Specifications (4-11), (4-12), (4-17), (4-18), (4-19), (4-20), and (4-21) take log linear forms from their corresponding theories while the forms of Equations (4-13), (4-14), (4-14), and (4-16) come from the Box-Cox analysis using pooled data. Estimation of models (4-19), (4-20), and (4-21) gave them no support. The complete stochastic inventory model explained only 17 percent of the variance in the dependent variable, as shown by the adjusted R-squared. The f test for inclusion of all regressors rejected their joint significance. Also, all independent variables had insignificant coefficients estimates as given by t tests. These two examinations reveal that the regressors are useless either in combination or separately as explanations of cash holdings. Also, estimation of Model (4-20) gave an insignificant estimate for the coefficient of lagged money balances. Thus, the complete stock adjustment inventory model reduces to Specification (4-18). Finally, none of the dummy coefficients differed from zero in Equation (4-21) reducing it to Equation (4-18). This result implies that new liabilities introduced by banks cannot be effectively modeled by dummy variables. These findings justify excluding Specifications (4-19), (4-20), and (4-21) from the empirical analysis that follows.

Before analyzing the remaining equations, the origin of the translog cost model, Equation (4-16), must be outlined. This specification restricts the parameters of input prices to equal the estimates of demand elasticities calculated from the translog cost model. The latter appear as $\bar{\beta}_2$, $\bar{\beta}_3$, and $\bar{\beta}_4$ in Model (4-16) and are formed from the constrained translog set of equations.

The Iterative Three-Stage Least Squares (I3SLS) elasticity estimates (see Appendix F) appear in Table VI and have significant magnitudes. The derived money demand elasticities imply the following. A negative value for own price elasticity means an inverse relationship between money balances and their price. The negative sign for η_{ml} implies that changes in labor's price cause money holdings to change in the opposite direction. In contrast, money substitutes for capital since η_{mk} exceeds zero. The inclusion of these elasticities in Equation (4-16) allows direct comparison of the translog cost model with the alternative specifications.

Table VI also contains computations of Allen partial elasticities. t tests show that all elasticities differ significantly from zero and implies following. Money compliments labor and substitutes for capital as given by the negative sign for σ_{ml} and positive sign of σ_{mk}. The former finding contradicts to the transactions cost argument that money balances should release labor services. The latter finding supports the neoclassical position that money holdings replace capital inventories. Finally, since σ_{lk} has a positive sign, labor and capital substitute for one another in production.

Obtaining these values requires estimation of the translog cost parameters which occurs after the following sequence of tests. First, the symmetry constraints were tested using pooled data and both I2SZ and I3SLS methods. Two f statistics result, one from each estimation method. In both cases, the restrictions are valid implying that the constrained form properly represents the translog cost model. The next step examines the disturbance characteristics of the I2SZ and I3SLS estimates of this equation.

The adjusted Durbin-Watson test results in a calculated value less

TABLE VI

TRANSLOG COST MODEL

Coefficient Parameter	I2SZ Estimate[1]	I3SLS Estimate	Macropa-rameter Estimate	Total Aggregation Bias[2]	Corresponding Bias	Noncorres-ponding Bias	True Parameter Value	Deviation from Expectation[3]	Elasticity Parameter	I3SLS Estimate	Separability Type	Test Statistic[4]
	Cost Share Equations								Elasticities		Separability	
q_{mm}	0.0566 (0.0079)	0.0584 (0.0054)	0.05617	0.00251	0.00123	0.00128	0.05032	0.00334	s_{mm}	-3.1600 (0.5416)	Global	10.4538*
q_{ml}	-0.1456 (0.0183)	-0.1372 (0.0116)	-0.15382	0.00027	-0.00302	0.00329	-0.15841	0.00432	s_{ml}	-0.9593 (0.1670)	Linear ML-K	8.56123*
q_{mk}	0.0649 (0.0203)	0.0788 (0.0171)	0.07614	0.00363	0.00158	0.00205	0.08031	-0.00780	s_{mk}	2.1250 (0.2443)	Linear MK-L	9.64186*
q_{lk}	-0.0284 (0.0193)	-0.0262 (0.0166)	-0.02556	0.00012	-0.00050	0.00062	-0.02930	0.00362	s_{ll}	-0.0952 (0.0099)	Linear LK-M	12.8523*
q_{ll}	0.1458 (0.0086)	0.1634 (0.0049)	0.15934	0.00641	0.00294	0.00347	0.14761	0.00532	s_{lk}	0.6258 (0.2366)	Non-Linear ML-K	6.18553*
q_{kk}	0.1536 (0.0711)	0.1372 (0.0556)	0.12959	0.00602	0.00316	0.00286	0.11437	0.00918	s_{kk}	-0.5713 (0.7947)	Non-Linear MK-L	7.78918*
w_m	0.6291 (0.1332)	0.6154 (0.1179)	0.62493	0.02748	0.01846	0.00902	0.60488	-0.00743	h_{mm}	-0.3160 (0.0541)	Non-Linear LK-M	5.93836*
w_l	0.8498 (0.2291)	0.8677 (0.2036)	0.86144	0.03904	0.02516	0.01388	0.83055	-0.00815	h_{ml}	-0.6715 (0.1670)		
w_k	0.4382 (0.1147)	0.4124 (0.0986)	0.38586	0.01856	0.01155	0.00701	0.37612	-0.00882	h_{mk}	0.9875 (0.1205)		

[1]Standard errors in parenthesis.

[2]See note 3, Table III.

[3]See note 4, Table III.

[4]Asterisk (*) designates acceptance of the null hypothesis at 1% significance level.

than the critical value for both I2SZ and I3SLS implying the presence of first order autocorrelation. Bartlett's test confirms the absence of heteroscedasticity in both cases since calculated values do not exceed critical values. The autocorrelation necessitates firm by firm generalized differencing as described previously.

The data, transformed by 192 individual firm autocorrelation coefficient estimates, was then subjected to I2SZ and I3SLS. Table VI contains the results which have homoscedastic and non-autoregressive disturbances. These properties imply that the estimates give an accurate basis for additional analysis. Furthermore, since no signficant difference exists between the estimates from the two methods, only those of I3SLS will be used in the analysis below. Since pooled data formed these estimates, the next task addresses the aggregation problem to validate the pooling technique.

A measure of the extent of aggregation bias appears in Table VI. Inspection shows total bias ranging from about three to six percent of the corresponding coefficient estimate magnitude. Whether or not these amounts constitute significant bias or not is a testable question.

The test for aggregation bias uses the f statistic formed by estimating unconstrained and constrained regressions. Since the calculated value does not exceed its critical magnitude, no significant aggregation bias exists. This result justifies the pooling of all observations and reaffirms the desirable properties of the I3SLS coefficient estimates. Therefore, they give an accurate basis for describing the production process as embodied in both Allen partial elasticities and derived factor demand elasticities.

These estimates also allow testing of the underlying production structures. Monotonicity of output with respect to all inputs exists since all fitted cost share estimates exceed zero. The production function has strictly convex isoquants because constructed bordered hessians are all negative definite. Finally, Table VI contains the results of all separability tests.

In every case, the exams support the absence of any type of weak separability and imply the following. Money balances must appear as an input in production function estimation. If not, biased estimates of the marginal products occur for all remaining inputs. For the same reason, estimating derived money demand must use a specification containing the prices of all other inputs. These implications support the neoclassical position and refute the transactions cost argument about money's role in the production function which the translog cost model embodies in the derived demand elasticities. Thus, Equation (4-16) allows direct comparisons of this model with the seven alternatives after discovering the superior estimation method.

Estimation

Ordinary Least Squares. Initially, the pooled data is subjected to ordinary least squares (OLS) for every model. Table VII contains the estimates of the eight specifications all of which have undesirable disturbance characteristics. The adjusted Durbin-Watson test indicates the presence of first order autocorrelation in every equation. Heteroscedasticity is absent from all models as shown by Bartlett's test. These findings necessitate firm by firm generalized differencing of every equation.

Generalized Difference Model. The transformed observations give

TABLE VII

ORDINARY LEAST SQUARES (ANNUAL)

Equation	Coefficient Estimate[1]	S	Box-Cox Test	Adjusted Durbin-Watson Test	Bartlett's Test
Equation (4-11) - TIM[2]		1.25426		0.35283	3.18245*
INT[3]	-2.5454				
	(0.3899)				
LN(S)	-0.9596				
	(0.0313)				
LN(FFR)	-0.3912				
	(0.1755)				
Equation (4-12) - DIM		1.17381		0.36614	3.05319*
INT	-2.7956				
	(0.4153)				
LN(S)	0.9752				
	(0.0345)				
LN(FFR)	-0.5958				
	(0.1820)				
Equation (4-13) - PM		1.06919	2.93418*	0.33939	2.86441*
INT	1.6210				
	(1.7862)				
LN(CA)	1.0636				
	(0.0276)				
LN(FFR)	-0.6536				
	(0.2290)				
LN(CBR)	-2.0618				
	(0.9589)				
LN(EQR)	-0.4959				
	(0.3614)				
LN(PPI)	-0.1740				
	(0.1394)				
Equation (4-14) - TLM		1.98731	3.48124*	0.22348	5.81214*
INT	-2.5379				
	(1.3299)				
LN(PPC)	0.3354				
	(0.3395)				
LN(FFR)	0.1259				
	(0.3703)				
LN(EQR)	0.1132				
	(0.5549)				
LN(WR)	2.8247				
	(0.8080)				
LN(RR)	0.3034				
	(0.2220)				
Equation (4-15) - NM		1.18219	3.18412*	0.42649	3.28432*
INT	-3.1863				
	(0.4947)				
LN(Q)	0.8950				
	(0.3150)				
LN(FFR)	-0.3026				
	(0.0976)				
LN(WR)	1.1094				
	(0.2946)				
LN(RR)	0.4661				
	(0.1294)				
Equation (4-16) - TRCM		1.15219	3.00541*	0.38123	2.98513*
INT	-2.1042				
	(0.3886)				
LN(Q)	0.6128				
	(0.1823)				
LN(FFR)	-0.3160				
LN(WR)	-0.6175				
LN(RR)	0.9875				
Equation (4-17) - CTIM		1.12555		0.47965	3.21862*
LN(RPV)	0.2556				
	(0.0953)				
LN(S)	0.8580				
	(0.0304)				
LN(FFR)	-0.6173				
	(0.1454)				
Equation (4-18) - CDIM		1.10948		0.48086	2.91574*
LN(RPV)	0.2543				
	(0.0934)				
LN(CR)	0.8584				
	(0.0291)				
LN(FFR)	-0.6265				
	(0.1257)				

[1]See note 1, Table II.

[2]All abbreviations correspond to those in Table I.

[3]Intercept term.

TABLE VIII

GENERALIZED DIFFERENCE MODEL (ANNUAL)

	Estimation					Aggregation							Multicollinearity			
													k=.20		k=.25	
Equation	Coefficient Estimate[1]	S	Box-Cox Test	Adjusted Durbin-Watson Test	Bartlett's Test	Macro-Parameter Estimate	Total Aggregation Bias[3]	Corresponding Bias	Non-Corresponding Bias	True Parameter Value	Deviation From Expectation[4]	F-Test	Coefficient Estimate	S	Coefficient Estimate	S
Equation (4-11) - TIM[2]		0.64613		2.04790*	2.33481*							1.0138*				
INT[5]	-3.4721 (0.3440)															
LN(S)	0.9429 (0.0423)					0.93179	0.01134	0.02115	-0.00981	0.91302	0.00743					
LN(FFR)	-0.2918 (0.0859)					-0.31058	0.00333	-0.00678	0.01011	-0.30579	-0.00812					
Equation (4-12) - DIM		0.63366		2.04249*	2.48166*							1.1255*				
INT	-3.2097 (0.3718)															
LN(CR)	0.8752 (0.0434)					0.86820	0.01101	0.01779	-0.00678	0.85327	0.00392					
LN(FFR)	-0.4471 (0.0914)					-0.46182	0.00581	-0.00851	0.01432	-0.46231	-0.00532					
Equation (4-13) - PM		0.57824	3.14861*	1.98609*	2.95414*							1.1483*				
INT	1.1867 (0.8889)															
LN(CA)	1.1147 (0.0400)					1.21349	0.00669	0.02675	-0.02006	1.19481	0.01199					
LN(FFR)	-0.6056 (0.1404)					-0.53492	-0.03936	-0.01090	-0.02846	-0.50142	0.00586		-0.5814	0.59613	-0.5387	0.60381
LN(CBR)	-2.3053 (0.5269)					-2.34266	-0.10374	-0.03458	-0.06916	-2.23179	-0.00713		-2.1688		-2.0034	
LN(EQR)	-0.4931 (0.2316)					-0.53001	-0.02565	-0.01036	-0.01529	-0.51221	0.00785		-0.4731		-0.4205	
LN(WPI)	-0.2344 (0.0931)					-0.23087	-0.01405	-0.00585	-0.00820	-0.20819	-0.00863		-0.2183		-0.1988	
Equation (4-14) - TCM		0.93998	2.83549*	1.89971*	2.86621*							1.3211*				
INT	-1.4251 (1.3282)															
LN(PPC)	0.5084 (0.3007)					0.48667	0.02389	0.01271	0.01118	0.46874	-0.00060					
LN(FFR)	0.1929 (0.1026)					0.17381	0.01225	0.00463	0.00762	0.15365	0.00791		0.1733	0.95511	0.1648	0.96863
LN(EQR)	0.1551 (0.4821)					0.17152	0.00760	0.00341	0.00419	0.16065	0.00327		0.1446		0.1287	
LN(WR)	2.2174 (0.5849)					2.20612	0.10643	0.04213	0.06430	2.11442	-0.01493		2.0654		1.9913	
LN(RR)	0.2052 (0.1436)					0.21334	0.01149	0.00613	0.00436	0.21007	-0.00822		0.1985		0.1743	
Equation (4-15 - NM		0.63818	2.14005*	2.16915*	3.59402*							1.0034*				
INT	-3.5307 (0.6570)															
LN(Q)	1.0118 (0.0586)					1.00812	0.03946	0.02327	0.01619	0.97512	-0.00646					
LN(FFR)	-0.2550 (0.0857)					-0.30119	0.00357	-0.00408	0.00765	-0.30014	-0.00542		-0.2481	0.64981	-0.2366	0.65014
LN(WR)	0.2710 (0.4813)					0.30425	0.01274	0.00569	0.00705	0.28019	0.01132		0.2603		0.2577	
LN(RR)	0.0707 (0.1200)					0.06528	0.00367	0.00170	0.00197	0.05881	0.00280		0.0661		0.0624	
Equation (4-16) - TRCM		0.62139	2.94813*	2.09811*	3.68829*							1.1186*				
INT	-3.0891 (0.5123)															
LN(Q)	0.8914 (0.0419)					0.88711	0.03399	0.02015	0.01384	0.86146	-0.00834					
LN(FFR)	-0.3160															
LN(WR)	-0.6715															
LN(RR)	0.9875															
Equation (4-17) - CTIM		0.62452		1.98789*	2.81424*							1.0889*				
LN(RPV)	0.1668 (0.0864)					0.18097	0.01062	0.00484	0.00578	0.17936	-0.00900					
LN(S)	0.6579 (0.0374)					0.65094	0.00856	0.01382	-0.00526	0.63218	0.01020					
LN(FFR)	-0.4263 (0.0933)					-0.46152	0.00298	-0.00810	0.01108	-0.44812	-0.00638					
Equation (4-18) - CDIM		0.60127		1.99644*	2.21377*							1.1356*				
LN(RPV)	0.2125 (0.0932)					0.25381	0.01216	0.00515	0.00701	0.23417	-0.00748					
LN(CR)	0.6120 (0.0397)					0.62279	0.00621	0.01180	-0.00559	0.60818	0.00840					
LN(FFR)	-0.4769 (0.1019)					-0.49018	0.00189	-0.01145	0.01334	-0.48612	-0.00595					

[1]See note 1, Table II.

[2]See note 2, Table VI.

[3]See note 3, Table III.

[4]See note 4, Table III.

[5]Intercept term.

the results in Table VIII, when subjected to OLS. Durbin-Watson and Bartlett's tests show that the disturbances of all models conform to the classical linear regression model assumptions. The estimates from the generalized difference model (GDM) are unbiased and efficient and, therefore, provide a basis for further analysis.

The initial step re-examines the functional forms of Equations (4-13), (4-14), (4-15), and (4-16). This repetition is desirable since the Box-Cox analysis bases the optimal form on the data which has undergone transformation. Using all transformed data pooled together, the analysis provides calculated statistics, for the test of a log linear form, appearing in Table VIII. In each case, the exam results in reconfirmation of the log-linear forms for the portfolio, transactions cost, neoclassical, and translog cost models. Thus, using pooled observations, all eight equations take linear in the natural logarithm specifications. Yet, the absence of aggregation bias is necessary to justify the pooling method. The extent of such bias must be measured and tested for significance.

Constructing the expected value of the macroparameters gives the calculations in Table VIII. The total bias (corresponding plus noncorresponding) ranged from two to six percent of the magnitude of the respective coefficient. The average amount of total bias for all variables equaled four percent implying a small inaccuracy due to aggregation. Moreover, f tests indicate that this amount is insignificant as shown by the f statistics in Table VIII. Hence, pooling can occur and the generalized difference model (GDM) renders unbiased and efficient estimates. However, in the presence of multicollinearity more accurate estimates may exist.

Table VIII contains the results of ridge regressions for Equations (4-13), (4-14), and (4-15) when k takes the values of .10 and .25, respectively. Since no substantial reduction occurs from OLS estimates, significant multicollinearity does not exist in the data bases. The GDM results provide unbiased and efficient estimates which are also precise and stable. Yet, superior estimates may result by adjusting for pooled data.

Pooling Adjusted Methods. Using the raw data (not transformed by GDM) the Error Components Model (ECM) gives estimates contained in Table IX. Two findings are germane. First, each model has a standard error of regression (S) reduced below that generated by OLS. The increase in explanatory power implies that intercept differences do exist. Given this property, the ECM estimators give more accurate results than those of OLS. Second, most coefficient estimates are smaller than OLS counterparts. Incorporating these cross-sectional and time series differences can account for the failure of previous literature to find significant economies of scale. While an improvement, the serial correlation structure suggests that these estimates do not represent the superior set.

The covariance model (COVM) generates more precise estimates, which appear in Table IX. Two results suggest the superiority of these estimates compared to those of the GDM. First, the f test for joint inclusion of all dummies supports the unconstrained regression (COVM) over that of a constrained regression (GDM) for each specification. Second, the lower standard error of regression (S) for the COVM equations relative to those from GDM gives further support to this conclusion. Consequently, the COVM will be employed for comparing the eight specifications.

TABLE IX

POOLING ADJUSTED MODELS (ANNUAL)

Equation	Covariance Model: Coefficient Estimate	S	Adjusted Durbin-Watson Test	Bartlett's Test	Zero Correlation Coefficient Test	Joint Inclusion f Test	Error Components Model: Coefficient Estimate	S
Equation (4-11) - TIM[2]		0.56949	2.14665*	3.81650*	0.1038*	831.021		0.62571
INT[3]	-1.3991				(0.1486)		-0.8571	
	(0.4403)						(0.5156)	
LN(S)	0.5391						0.6367	
	(0.0923)						(0.0689)	
LN(FFR)	-0.2454						-0.3625	
	(0.0764)						(0.1413)	
Equation (4-12) - DIM		0.56495	2.15895*	3.66712*	0.0986*	835.774		0.62079
INT	-0.7842				(0.1337)		-1.4250	
	(0.4457)						(0.5614)	
LN(CR)	0.5474						0.7310	
	(0.0870)						(0.0745)	
LN(FFR)	-0.3923						-0.4331	
	(0.0801)						(0.1526)	
Equation (4-13) - PM		0.52339	2.14123*	3.28164*		840.903		0.55066
INT	0.9130						1.2902	
	(0.8751)						(1.7824)	
LN(CA)	1.1457						1.1553	
	(0.1144)						(0.0702)	
LN(FFR)	-0.5311						-0.6882	
	(0.1250)						(0.2273)	
LN(CBR)	-2.2825						-1.0841	
	(0.4679)						(1.1150)	
LN(EQR)	-0.2022						-0.6068	
	(0.2806)						(0.0365)	
LN(RPI)	-0.0912						-0.1709	
	(0.1831)						(0.1375)	
Equation (4-14) - TCM		0.82692	2.06391*	4.03812*	0.4814	601.452		0.87071
INT	-3.8656				(0.1689)		3.8611	
	(1.2934)						(1.0877)	
LN(PPC)	0.5382						-0.0803	
	(0.1931)						(0.1905)	
LN(FFR)	-0.1759						0.0690	
	(0.0794)						(0.2447)	
LN(EQR)	0.1824						0.2824	
	(0.3872)						(0.3481)	
LN(WR)	1.6466						-0.6228	
	(0.5974)						(0.7140)	
LN(RR)	0.1280						-0.0983	
	(0.1196)						(0.1027)	
Equation (4-15) - NM		0.56312	2.20434*	3.46812*	0.1007*	793.664		0.62815
INT	-0.9142				(0.1438)		-1.9818	
	(0.8984)						(0.9200)	
LN(Q)	0.3829						0.7582	
	(0.1176)						(0.0809)	
LN(FFR)	-0.2843						-0.3553	
	(0.0775)						(0.1617)	
LN(WR)	-1.2181						-0.2181	
	(0.5532)						(0.5422)	
LN(RR)	0.0322						-0.1298	
	(0.1192)						(0.0976)	
Equation (4-16) - TRCM		0.55220	2.10552*	3.00412*	0.1099*	811.439		0.61021
INT	-0.7788				(0.1387)		-1.8864	
	(0.1513)						(0.7122)	
LN(Q)	0.4866						0.8341	
	(0.0889)						(0.0724)	
LN(FFR)	-0.3160						-0.3160	
LN(WR)	-0.6175						-0.6175	
LN(RR)	0.9875						0.9875	
Equation (4-17) - CTIM		0.55389	2.21300*	3.84913*	0.0983*	822.588		0.61931
LN(RPV)	0.3604				(0.1254)		0.2113	
	(0.0600)						(0.0579)	
LN(S)	0.4475						0.7881	
	(0.0456)						(0.0805)	
LN(FFR)	-0.4009						-0.4880	
	(0.0908)						(0.1377)	
Equation (4-18) - CDIM		0.54274	2.21135*	3.77312*	0.0979*	837.166		0.61753
LN(RPV)	0.3889				(0.1301)		0.2320	
	(0.0801)						(0.0785)	
LN(CR)	0.4450						0.8298	
	(0.0466)						(0.0824)	
LN(FFR)	-0.4064						-0.4791	
	(0.0913)						(0.1559)	

[1]See note 1, Table II.

[2]See note 2, Table VI.

[3]Intercept term.

Evaluation Based on Estimation. The criteria used in evaluating the estimated models include: magnitude of S^2 (estimated specification variance), joint significance tests of all regressors, individual significance tests for each regressor, and signs and magnitudes of coefficient estimates.

The magnitude of S^2 measures the specification's inaccuracy in fitting values for the actual outcomes of the dependent variable. The superior specification will have a significantly lower S^2 than that of the alternative models. Inspection of the S^2 values, appearing in Table IX, renders the ranking of models from most to least accurate as: portfolio, complete deterministic inventory, translog cost, complete traditional inventory, neoclassical, deterministic inventory, traditional inventory, and transactions cost. Testing for significant differences requires using the zero correlation coefficient procedure. Table IX contains values for r between the portfolio model and each of the seven alternatives along with standard errors. Only the transactions cost specification has inferior explanatory power. No statistical difference exists between the performance of the seven remaining models.

While important, these tests only evaluate the relative ability of the models. Some interest also lies in their absolute explanatory power. The f test for joint inclusion of all explanatory variables gives an absolute criterion. Inspection of Table IX shows that the f value exceeds the critical value for all eight models. Every specification contains regressors which, taken together, significantly influence the dependent variable.

The next test examines the explanatory power of individual regressors. The t tests for significance contained in Table XI show that all regressors of the four inventory models and the translog cost model have significant magnitudes. The portfolio specification includes two insignificant regressors: equity returns (EQR) and the percentage change in the product price index (PPI).[5] In the transactions cost model, both EQR and the price of capital (RR) have no influence. Finally, RR has no explanatory power within the neoclassical specification. These results raise the four inventory models and the translog cost model above their competitors.

Further distinctions come from the final estimation criteria of a coefficient's sign and magnitude. All four inventory models have coefficient estimates with correct signs. Their magnitudes, as judged by t tests, give the following results. The coefficient for the transactions proxy equals one-half in all cases. However, the traditional inventory model, Equation (4-11), has an interest rate elasticity different from a negative one-half while the transfer cost proxy in the complete traditional inventory model, Equation (4-17), differs from one-half. In contrast, all estimates of the deterministic inventory, model (4-12), and complete deterministic inventory, model (4-18), equations equal their hypothesized values. These findings suggest the following conclusions. The two deterministic inventory specifications (using cash receipts) rank above the two traditional inventory models (using sales). Since the complete deterministic inventory model represents a general version of the deterministic inventory model, supporting the former simultaneously rejects the latter. Support exists not only for cash receipts as the transactions proxy but for economies of scale in money balances.

---

[5]The insignificant coefficient for PPI substantiates the homegeneous property of the portfolio equation.

Finally, such holdings vary directly with cash receipts and the inverse of repurchase agreement volume and indirectly with the federal funds rate.

The translog cost model's only estimated coefficient takes the correct sign. Since this output elasticity does not differ from one-half, it implies economies of scale in cash holdings. The remaining (non-estimated) coefficient values imply that cash balances vary indirectly with the federal funds rate and wage rates but change directly with the user cost of capital.

The portfolio model gets correct signs for all significant coefficient estimates. The elasticity of current assets exceeds one, a denial of economies of scale. Both the federal funds rate and the corporate bond rate have an inverse effect on money holdings.

The transactions cost model also shows correct signs for all of the significant coefficient estimates. Money balances change in the same direction as the price index for commodities and wage rates. The federal funds rate has an inverse affect on cash holdings.

Finally, the neoclassical model's coefficient estimates all take the correct signs. Changes in output alter money holdings in the same direction. Cash balances vary inversely with both the wage rate and the federal funds rate.

In summary, the complete deterministic inventory and the translog cost models both receive full support from estimation criteria. The remaining specifications each fail at least one evaluation. No further distinction occurs between the two fully supported models based on estimation measures. Thus, prediction criteria must make any additional ranking of models.

Prediction

Formation of a Forecast Series. Predictions from the eight specifications use estimates from the COVM adjusted for serial correlation. The actual outcomes to be predicted by the models exist for every firm in each year 1979 and 1980, giving 384 total observations. Each specification makes one step ahead forecasts of the 384 values which results in a forecast series.

Evaluation of the Forecast Series. Evaluation of the forecast series rests on both absolute and relative prediction criteria. Absolute measures included mean square error (MSE) and simple regressions. Relative criteria come from composite predictors.

Absolute Criteria. Calculations of MSE for each specification appear in Table X. Ordering their magnitudes from smallest to largest gives the ranking for models as: the portfolio, the complete deterministic inventory, the translog cost, the complete traditional inventory, the neoclassical, the deterministic inventory, the traditional inventory, and the transactions cost. Asserting the superiority of the portfolio model requires that a significant difference exists between its MSE and that of alternative specifications. The test for zero correlation coefficient can make this determination. The value of r, and its standard error, for the portfolio model against each of the seven alternatives appears in Table X. Test results show that the transactions cost model makes inferior predictions and implies no significant difference in the accuracy of the remaining seven equations. While the MSE provides a measure of relative accuracy, a specification should make optimal forecasts in an absolute sense.

A predictor series is optimal if its error series has zero mean and no autocorrelation. The t test for significance can evaluate the mean of the error series. Table X contains the examinations which indicate that each of the eight specifications makes unbiased forecasts. The adjusted Von-Neuman ratio test can determine the autocorrelation of the error series. The statistic values appear in Table X and show that first order autocorrelation does not exist in any equation. Combining this property with the finding of unbiasedness means that all eight specifications make optimal forecasts.

To double check these conclusions, optimality can be tested using the predictor series itself. The regression of predicted values (P) on actual outcomes (M) as,

$$M_t = \alpha_0 + \alpha_1 P_t + u_t \tag{4-22}$$

where u is a disturbance term indicates an optimal forecaster when estimates of α_0 and α_1 do not significantly differ from zero and one, respectively. Standard t tests can evaluate these magnitudes. The regression results appear in Table X. In all eight cases, the estimates of α_0 and α_1 do not differ from zero and one, enforcing the existence of optimal properties for the eight predictors. Even though all specifications have desirable properties, their prediction accuracy may differ.

A measure of precision can come from the regression results. From a set of optimal predictors, the most accurate has the smallest S^2 calculated from the simple regressions. Inspection of S^2 magnitudes results in ranking the models from most to least precise as: the portfolio, the complete deterministic inventory, the translog cost, the complete traditional inventory, the deterministic inventory, the

TABLE X

PREDICTION RESULTS (ANNUAL)

Model	Error Series: MSE	Error Series: Test for Zero Correlation Coefficient[1]	Error Series: Test for Zero Mean	Error Series: Non-Nueman Ratio	Predictor Series: Simple Regressions: INT	Predictor Series: Simple Regressions: P_1	Predictor Series: Simple Regressions: S	Predictor Series: Simple Regressions: Test for Zero Correlation Coefficient	Composite Predictors Unconstrained Then Constrained[3]: P_1 P_2 P_3 P_4 P_5 P_6 P_7 and P_8	P_3 and P_1	P_3 and P_2	P_3 and P_4	P_3 and P_5	P_3 and P_6	P_3 and P_7	P_3 and P_8
Model (4-11) TIM[2]	0.59810	0.1052 (0.1837)	0.1081 (0.2804)	2.00367	0.3196 (0.3007)	0.8861 (0.2182)	0.56614	0.0993 (0.1548)	0.1401 (0.1277)	0.2861 (0.2104)						
		C=2.58		C=1.44				C=2.58	0.1318 (0.1076)	0.0782 (0.1084)						
Model (4-12) DIM	0.59333	0.1018 (0.1551)	0.0923 (0.2415)	2.10991	0.2958 (0.3452)	0.9025 (0.2014	0.54812	0.0834 (0.1498)	0.0834 (0.1045)		0.2673 (0.1955)					
		C=2.58		C=1.44				C=2.58	0.0756 (0.0954)		0.0711 (0.0976)					
Model (4-13) PM	0.52869		0.0845 (0.1760)	1.98915	0.1889 (0.2016)	0.9308 (0.0907)	0.51814		0.1531 (0.0961)	0.8875 (0.1192)	0.8912 (0.1153)	0.7954 (0.1811)	0.9068 (0.0971)	0.9114 (0.1025)	0.9351 (0.1240)	0.9412 (0.1176)
				C=1.44					0.1029 (0.0557)	0.9218 (0.1084)	0.9289 (0.0976)	0.8057 (0.1745)	0.9441 (0.0914)	0.9489 (0.1004)	0.9577 (0.1125)	0.9604 (0.1079)
Model (4-14) TCM	0.88027	0.6086 (0.1947)	0.2313 (0.4055)	2.21147	0.3868 (0.2711)	0.8663 (0.3119)	0.68160	0.5535 (0.1782)	0.3179 (0.2837)			0.3412 (0.2555)				
		C=2.58		C=1.44				C=2.58	0.2664 (0.2538)			0.1943 (0.1745)				
Model (4-15) NM	0.59140	0.1197 (0.1634)	0.0961 (0.2556)	1.99185	0.3267 (0.3099)	0.8966 (0.1214)	0.54987	0.1044 (0.1496)	0.1066 (0.1454)				0.2755 (0.1882)			
		C=2.58		C=1.44				C=2.58	0.0953 (0.1287)				0.0559 (0.0914)			
Model (4-16) TRCM	0.57994	0.1189 (0.1832)	0.0884 (0.1956)	2.04831	0.2840 (0.3147)	0.9291 (0.0978)	0.53315	0.0883 (0.1357)	0.1168 (0.1375)					0.2511 (0.1926)		
		C=2.58		C=1.44				C=2.58	0.0992 (0.1401)					0.0511 (0.1004)		
Model (4-17) CTIM	0.58172	0.1018 (0.1486)	0.0903 (0.2510)	2.00396	0.3042 (0.2996)	0.9104 (0.1276)	0.53861	0.0942 (0.1228)	0.2010 (0.1281)						0.1667 (0.2193)	
		C=2.58		C=1.44				C=2.58	0.1252 (0.1414)						0.0423 (0.1125)	
Model (4-18) CDIM	0.57001	0.0992 (0.1225)	0.0859 (0.2014)	1.98932	0.2723 (0.2805)	0.9245 (0.1073)	0.52832	0.0766 (0.1145)	0.0991 (0.1004)							0.1784 (0.1921)
		C=2.58		C=1.44				C=2.58	0.1036 (0.0952)							0.0396 (0.1079)
S[4]									0.44321	0.50184	0.51004	0.52813	0.50456	0.49872	0.49014	0.50061
									0.45607	0.52366	0.53630	0.53994	0.51137	0.50914	0.50336	0.51462

[1]Standard errors in parenthesis: Asterisk (*) designates acceptance of null hypothesis at 1% significance level.

[2]See note 2, Table VI.

[3]See note 2, Table V.

[4]See note 3, Table V.

neoclassical, the traditional inventory, and the transactions cost model. While differences exist, they may not have significant magnitudes. The zero correlation coefficient test can determine which, if any, models give superior performance.

The computed values of r and its standard error for each specification with the portfolio model appear in Table X. Only the transactions cost equation gives inferior performance. The remaining seven specifications make forecasts, which are not significantly different.

In summary, criteria based on the error series gives identical evaluations as criteria from the predictor series. All eight specifications make optimal forecasts. The transactions cost model is less accurate than the remaining seven models. Finally, predictions from these seven specifications are statically equal. Absolute prediction criteria make no distinction between Equations (4-11), (4-12), (4-13), (4-15), (4-16), (4-17), and (4-18). The final evaluation comes from relative forecast performance within a composite predictor.

Relative Criteria. The composite predictor containing all eight specifications takes the form,

$$M = \alpha_1 P_1 + \alpha_2 P_2 + \alpha_3 P_3 + \alpha_4 P_4 + \alpha_5 P_5 + \alpha_6 P_6 + \alpha_7 P_7 + \alpha_8 P_8 + \varepsilon \quad (4\text{-}23)$$

where M represents actual outcomes, P_i stands for the prediction series from the i^{th} specification, and ε is a disturbance term. When all forecasts make unbiased predictions, the α_i must sum to unity.[6] Estimates

---

[6]Testing this characteristic necessitates estimating Equation (4-23) and a constrained regression (which assumes the α_i sum to unity) equal to,

for Equation (4-23) and its constrained form appear in Table X. The f statistic exceeds the critical value, implying that the coefficient estimates sum to unity. This conclusion supports the unbiased property of the eight specifications. Furthermore, the t values show that all regressor coefficients in both equations are insignificant. None of the specifications make a marginally significant contribution to the remaining group. Yet, these t statistics may be deflated due to multicollinearity. Avoiding this problem necessitates construction of an alternative set of composite predictors.

Seven regressions are estimated, each taking the form,

$$M = \alpha_3 P_3 + \alpha_i P_i + \varepsilon \tag{4-25}$$

where i = 1, 2, 4, 5, 6, 7, 8. When all predictors are unbiased, the coefficients of each regression sum to unity. Testing this property requires estimating seven constrained regressions as,

$$e_3 = \alpha_i (p_i - P_3) + \varepsilon. \tag{4-26}$$

Another examination uses t tests to measure the marginal contribution of the i^{th} predictor. If P_3 contains all the useful information of P_i, then estimates of α_i and α_3 from Equation (4-25) will not differ from

---

[6](Continued)

$$e_3 = \alpha_1(P_1 - P_3) + \alpha_2(P_2 - P_3) + \alpha_4(P_4 - P_3) + \alpha_5(P_5 - P_3) + \alpha_6(P_6 - P_3)$$
$$+ \alpha_7(P_7 - P_3) + \alpha_8(P_8 - P_3) + \varepsilon \tag{4-24}$$

where e_3 represents the error series from the portfolio model. The results of estimation can construct an f statistic. When its value exceeds the critical value, the test upholds the restrictions.

zero and one. With respect to Equation (4-26) when P_i has no information to add to P_3, the estimates of α_i will not differ from zero.

The results of all 14 regressions appear in Table X. For all eight f tests, the calculated value was less than the critical value. This finding reaffirms the property of lack of bias in these predictors. In unconstrained form, the seven regressions had estimates of α_i and α_3 not different from zero and unity, respectively. P_3 contains all the useful information in the alternative forecasts. Using the seven constrained regressions, none of the α_i differed from zero which confirms the prior finding.

In summary, no distinctions between the models come from relative prediction criteria. None make a marginally significant contribution to the remaining seven models. Absolute prediction criteria indicate that all models make optimal forecasts. The transactions cost specification gives less accurate predictions than the seven alternatives. Finally, these seven make forecasts which do not differ significantly in terms of accuracy.

Summary. Combining the findings from both prediction and estimation criteria renders the following. The complete deterministic inventory and the translog cost models pass every estimation examination. The others fail on at least one account. Since these two models predict as accurately as the remaining specifications, they provide superior specifications of firm's behavior.

Summary

This chapter has conducted a systematic econometric analysis of firms' money demand. Using the Box-Cox analysis when necessary, specific comparable equations reflecting the unique aspects of each money demand theory were formed. For both data bases, the superior econometric technique applies the covariance model to all observations transformed by generalized differencing. The resulting estimates of and predictions from each specification provide for model evaluation based on performance criteria.

Judging the findings in light of these criteria renders the following conclusions. The deterministic inventory models (using cash receipts) pass every estimation exam. Since they also explain and predict at least as well as the alternatives, inventory theory provides the superior specifications of money demand. The estimates imply that cash holdings exhibit economies of scale with respect to transactions, vary inversely with their opportunity cost, and change in the same direction as transfer costs. Cash receipts represent a more useful transactions variable than sales. Finally, the inverse volume of repurchase agreements gives an accurate measure for transfer costs while the federal funds rate properly reflects money's opportunity cost.

The other two theories give specifications that fall short of the deterministic inventory equations. Portfolio theory has an model containing insignificant coefficient estimates. Thus, some of the postulated regressors add no useful information about money holdings. Such a result places the portfolio equation below inventory models even though they predict and explain with equal accuracy. Production theory gives mixed results. The transactions cost model not only has insignificant regressors, the significant coefficients take the wrong sign in some cases. These results, coupled with less explanatory and predictive power, indicate the inferiority of the transactions cost equation. The neoclassical specification also renders coefficient estimates with insignificant magnitudes. This result leaves it behind the inventory models even though it has equal explanatory and predictive power. Finally, while the translog cost equation passes all estimation tests, it does so only for the annual data base. Although this model gives the best representation of production theory, its performance relative to alternative theories remains incomplete due to lack of quarterly data.

CHAPTER V

SUMMARY AND CONCLUSIONS

The dissertation has selected inventory theory as the most useful model of firms' money demand. Since a theory's usefulness depends on its ability to explain and predict behavior, the choice employed both estimation and prediction criteria. They evaluated the performance of specifications from three different approaches to money demand: inventory models, a portfolio model, and production models. Previous theoretical work in each area culminates in the basis for testable equations embodying the unique characteristics of each theory. Completing the task in a manner that allows useful comparisons necessitated applying answers to theoretical questions uniformly to all three models. Since the resulting final specifications describe individual firm behavior, evaluation required individual firm observations. Combining these data with the most powerful econometric techniques rendered statistics measuring the explanatory and predicitve performance of each theory. The three parts of this methodology--specifications, data, and techniques--stem from analysis of existing empirical work on firms' money demand. Each is recapped below.

Specifications were built on analysis concerning several topics relevant to money demand functions. Previous empirical work has not approached these issues systematically and, thus, they provide no consensus results about money demand. The major points made in the dissertation pertaining to these topics are restated below.

Money takes a narrow definition as currency plus demand deposits. This restrictive view stresses money's unique asset characteristics: it acts as a medium of exchange and entails no risk of either default or loss in nominal value. Focusing on these attributes suggests comparing money to its alternatives in the same terms. A liquidity hierarchy emerges where assets are ranked with respect to their moniness. Those associated with greater transfer costs and/or risk lie farther from cash on the spectrum. They render attributes (including sizable returns) unlike those available from holding money. The asset closest to money would have zero transfer costs and entail identical riskiness and, thus, have qualities almost identical to money. The hierarchy not only shows how alternatives relate to one another, it provides a solution to another problem.

Money's opportunity cost equals the return on its best alternative use. If cash vanished, investors would hold the asset most like it to acquire money's two unique attributes. Thus, the price of money equals the return on a transfer costless asset with the same risk as money. While checkable deposits fill this role, government price ceilings make their return an unfit indicator of money's price. Fortunately, another near money, repurchase agreements, has a return established by free market forces and, therefore, capable of accurately reflecting the opportunity cost of money. While data on this return is not readily available, bank arbitrage ensures that it moves with the federal funds rate. This latter return provides an accurate measure of money's opportunity cost. In addition to alterations of all models suggested by these general topics, certain issues adjusted to the specifications of specific models.

Inventory theory postulates economies of scale in cash holdings which implies a distinctive relationship between firms. Economies mean that average transactions costs (ATRC) fall monotonically as transactions levels increase. Firms have incentive to expand or merge until the decline in ATRC matches the increase in average production costs caused by bureaucratic inefficiences. Alternatively, firms' can consolidate transactions in one institution. The bank will specialize in performing the middleman role between lenders and borrowers of funds. To provide this service, they must hire money from firms which requires a payment equal to money's opportunity cost. Because of price ceilings on demand deposits, banks make this payment in the form of services and by assuming the transactions function of firms. These payments induce businesses to voluntarily hold demand deposits in excess of any compensating balance. An involuntarily held compensating balance would constitute irrational behavior when reserve requirements exist. Thus, inventory theory is not invalidated by such an arrangement.

This basic economic relationship between firms and banks has broadened due to the expanding use of cash management techniques. Banks have marketed repurchase agreements in response to losing traditional deposits in an era of rising interest rates on and falling transfer costs of non-bank alternatives. Price ceilings prevented banks from making traditional deposits competitive. To reattract lost funds, they introduced a new liability with unique characteristics. Thus, an increase in RP volume should accompany the decline in transfer costs of non-bank alternatives as banks fight back. This analysis implies that the inverse of repurchase agreement volume gives an accurate measure of transfer costs. Combining these analyses with previous work on inventory theory renders several testable specifications.

With regard to portfolio theory, conclusions about specific issues include the following. The general equation from Friedman (1956) must be amended in two ways: the opportunity cost of money should appear and the general price level must be replaced with a product price index. The Box-Cox analysis selects a log linear specification for this amended form.

Finally, one issue pertains to production theory as follows. Since money demand derives from product demand, its form depends on the functions assumed in the profit maximizing process. The transactions cost variant employs a physical input production function coupled with a money requirements function. In contrast, the neoclassical version postulates a money balance inclusion production function. In both cases, a general form for the money demand results which the Box-Cox analysis puts into a log linear specification. Finally, the translog cost variant postulates a specific dual cost function that renders a testable specific form.

The analysis on these topics establishes equations which highlight each theory's unique approach to firms' money demands. Evaluating their performance as explanations and predictors of individual firm behavior requires incorporating the two remaining pieces of the methodology.

Since each equation gives a microfoundation, a complete test of their ability requires individual firm observations. Two samples of such data appear in the dissertation: one measured quarterly and another containing annual observations. They allow construction of the aggregation bias inherent in the estimates of macroparameters. Since this bias existed in insignificant amounts for both samples, pooling the data became an efficient strategy. This grouping necessitated the final part of the methodology.

Efficiency gains can occur in the estimators when using econometric techniques more powerful than ordinary least squares. Applying the error components model substantiated this hypothesis. Yet, the covariance model gave the most efficient unbiased estimators of all the available techniques. The corresponding estimates and predictions best reflect the usefulness of each theory. Thus, the criteria used them to make the model selection.

Applying the methodology renders several important conclusions, the most startling being the nearly identical performance from all models except the transactions cost equation. Despite the large amount of data variation the remaining production equations', the portfolio model's, and inventory specifications' overall explanatory and predictive performances are not statistically different. Distinctions are not made until applying the criteria of coefficients' signs and magnitudes. They separate out deterministic inventory theory as the most useful description of firms' behavior. The implications for money demand were as follows.

The adjusted R-square exceeded 90 percent in both samples. Cash balances exhibited economies of scale with respect to transactions. The magnitude of the transactions elasticity did not differ from one-half. Money holdings vary inversely with their own price and directly with transfer costs. Their elasticity estimates equaled the hypothesized values of a negative one-half and one-half, repectively.

To illustrate the importance of combining pooled individual firm data and advanced econometric techniques, the ordinary least squares results appear below. The adjusted R-square was less than 75 percent which understates the true explanatory ability of an inventory equation. The transactions elasticity equaled one in denial of economies of scale. While own price and transfer costs helped explain cash balances, their elasticity estimates differed from hypothesized values. The contrast of these results with those from the COVM illustrate the importance of pooling firm level data and using advanced econometric methods.

The COVM findings for inventory models compare to previous empirical work as follows. The high degree of explanatory power contrasts with the poor performance found by Sprenkle (1969). In fact, the R-squared values exceed those of any other inventory model study. The finding of economies of scale in transactions parallels that of Ben Zion (1974) and Vogel and Maddala (1967). Results of studies by Whalen (1965) and Meltzer (1963a) refute this property. The interest rate elasticity takes the same magnitude as that found by Ben Zion (1974). Finally, the transfer cost coefficient has no counterpart in previous studies.

The COVM estimates for the portolio equation explain a greater portion of variation in cash balances than any previous portfolio study. The results indicate the absence of economies of scale with respect to wealth. While Meltzer (1963a) finds no economies of scale with respect to sales, the work of Whalen (1965) supports economies with respect to sales as a percent of wealth. Neither of these results compare directly to this dissertation's finding. Also, the estimates of rates of return have no predecessors from previous work.

The neoclassical model renders COVM estimates which give the following comparisons with previous work. The overall explanatory power of this dissertation's estimates equals that of the study by Nadiri (1969).

Both papers imply economies of scale, negative own price elasticity, and substitution between labor and money. While Nadiri (1969) shows that capital substitutes for money, this study finds no relationship between the two. Finally, the strength of the substitutability between labor and money is greater here than in the Nadiri (1969) paper.

The COVM estimates for the translog cost model give the following comparisons. They support placing money into the production function. The same conclusion was reached by Sinai and Stokes (1972), Khan and Kouri (1975), Short (1979), and Dennis and Smith (1978). The money demand elasticities give comparisons to those of Dennis and Smith (1978). Both studies found negative values for own price elasticities and positive magnitudes for the cross elasticity of capital with respect to money. This study rendered a much larger value for the latter figure. Also, the negative value for the cross elasticity of labor with respect to money contrasts with their finding of a positive magnitude. While these results add new light on firms' money demand, future work should continue in three areas.

First, adequate estimation of the stochastic inventory variants has not occurred. Such a task awaits widely available, precise data on daily cash flow variances. The dissertation's crude attempt to measure this varible with its quarterly and annual counterparts failed to capture any effect on money balances.

Second, a more comprehensive comparison between inventory equations and a translog cost model should be undertaken. A quarterly data base would greatly enhance the initial comparison conducted in the dissertation. Currently, the variables needed to construct the translog model are not measured quarterly.

Finally, the methodology developed in the dissertation should extend to a systematic evaluation of household cash holdings. Individuals could adhere to either an inventory or portfolio model. If inventory theory proved more useful, a synthesis of household and firm sector money demand could occur. In contrast, a two-part additive function might result if the portfolio approach best described individual behavior. In either case, a total private sector money demand function could be formed which would accurately explain and predict money balances. Coupled with control over the money supply, it would promote social welfare by aiding in the formation of intelligent policy decisions.

A final comment should stress that sometimes simple is best. Model building represents a classic example of tradeoffs. A theory must include enough explanatory variables to accurately describe reality. Yet, it must filter out minor influences to remain manageable and applicable to many cases. Inventory theory filters more variables than the two alternatives and yet performs at least as well. This result comes very close to being a free lunch, something which should never be refused.

A SELECTED BIBLIOGRAPHY

Allen, R. G. (1965). Mathematical Economics. Second Edition. London: Macmillan and Co., Ltd.

Barro, R. (1970). "Inflation, the Payments Period, and the Demand for Money." Journal of Political Economy, Vol. 78, 1228-1263.

_______________ (1976). "Integral Constraints and Aggregation in an Inventory Model of Money Demand." Journal of Finance, Vol. 31, 77-88.

Barro, R., and S. Fischer (1978). "Recent Developments in Monetary Theory." Journal of Monetary Economics, Vol. 4, 133-167.

Barro, R., and A. Santomero (1972). "Household Money Holdings and the Demand Deposit Rate." Journal of Money Credit and Banking, Vol. 4, 397-413.

Bates, J. M., and C. W. Granger (1969). "The Combination of Forecasts." Operations Research Quarterly, Vol. 20, 451-468.

Baumol, W. (1952). "The Transactions Demand for Cash: An Inventory Theoretic Approach." Quarterly Journal of Economics, Vol. 66, 545-556.

Ben-Zion, U. (1974). "The Cost of Capital and the Demand for Money by Firms." Journal of Money, Credit, and Banking, Vol. 6, 838-845.

Ben Zion, U., and V. Ruttan (1975). "Money in the Production Function: An Interpretation of Empirical Results." Review of Economics and Statistics, Vol. 57, 246-247.

Berndt, E. R., and L. R. Christensen (1973a). "The Translog Function and the Substitution of Equipment, Structures, and Labor in U. S. Manufacturing, 1929-1968." Journal of Econometrics, Vol. 1, 81-113.

_______________ (1973b). "The Internal Structure of Functional Relationships: Separability, Substitution, and Aggregation." Review of Economic Studies, Vol. 40, 403-410.

_______________ (1974). "Testing for the Existence of a Consistent Aggregate Index of Labor Inputs." American Economic Review, Vol. 64, 391-404.

Berndt, E. R., and D. O. Wood (1975). "Technology, Prices, and the Derived Demand for Energy." The Review of Economics and Statistics, Vol. 52, 259-268.

Binswanger, H. P. (1974). "A Cost Function Approach to the Measurement of Elasticities of Factor Demand and Elasticities of Substitution." American Journal of Agricultural Economics, Vol. 56, 377-384.

Blackorby, C., D. Primont, and R. R. Russell (1977). "On Testing Separability Restrictions With Flexible Functional Forms." Journal of Econometrics, Vol. 5, 195-209.

_________________ (1978). Duality, Separability, and Functional Structure: Theory and Empirical Applications. New York: American Elsevier.

Blackorby, C., and R. R. Russell (1976). "Functional Structures and the Allen Partial Elasticities of Substitution: An Application of Duality Theory." Review of Economic Studies, Vol. 43, 285-292.

Box, G. E. P., and D. P. Cox (1964). "An Analysis of Transformations." Journal of the Royal Statistical Society, Series B, Vol. 26, 211-252.

Chetty, V. K. (1968). "Pooling of Time Series and Cross Section Data." Econometrica, Vol. 36, 279-290.

Christensen, L. R., and W. H. Greene (1976). "Economies of Scale In U. S. Electric Power Generation." Journal of Political Economy, Vol. 84, 655-676.

Christensen, L. R., and D. W. Jorgenson (1970). "U. S. Real Product and Factor Input, 1929-1967." Review of Income and Wealth, Series 16, 19-50.

Christensen, L. R., D. W. Jorgenson, and L. J. Lau (1973). "Transcendental Logarithmic Production Frontiers." Review of Economics and Statistics, Vol. 55, 28-45.

Cochrane, D., and G. H. Orcutt (1949). "Application of Least Squares Regression to Relationships Containing Autocorrelated Error Terms." Journal of the American Statistical Association, Vol. 44, 32-61.

Cooper, J. P., and C. R. Nelson (1975). "The Ex Ante Prediction Performance of the St. Louis and F.R.B.-M.I.T.-Penn. Econometric Models and Some Results on Composite Predictors." Journal of Money, Credit, and Banking, Vol. 7, 1-32.

Dennis, E., and V. Smith (1978). "A Neoclassical Analysis of the Demand for Real Cash Balances by Firms." Journal of Political Economy, Vol. 86, 793-811.

Denny, M., and M. Fuss (1977). "The Use of Approximation Analysis to Test for Separability and the Existence of Consistent Aggregates." American Economic Review, Vol. 67, 404-418.

Dhrymes, P. J. (1970). "Equivalence of Iterative Aitken and Maximum Likelihood Estimates for a System of Regression Equations." Econometrica, Vol. 43, 281-298.

_________________ (1972). "Criteria for Evaluation of Econometric Models." Annals of Economic and Social Measurements, Vol. 1, 291-324.

Diewart, W. E. (1974). "Applications of Duality Theory." In M. D. Intriligator and D. A. Kendrick, Eds., Frontiers of Quantitative Economics. Vol. II. Amsterdam: North-Holland Publishing Company, 106-171.

Dow Jones News Service. New York: Dow, Jones and Company, 1960-1981.

Draper, N. R., and C. V. Nostrand (1979). "Ridge Regression and James-Stein Estimation: Review and Comments." Technometrics, Vol. 21, 451-464.

Draper, N. R., and H. Smith (1966). Applied Regression Analysis. New York: John Wiley and Sons, Inc.

Enzler, J., L. Johnson, and J. Paulus (1976). "Some Problems of Money and Demand." Brookings Papers on Economic Activity, Vol. 3, 261-280.

Eppen, G., and E. Famma (1969). "Cash Balance and Simple Dynamic Portfolio Problems With Proportional Costs." International Economic Review, Vol. 12, 119-133.

Feige, E., and M. Parkin (1971). "The Optimum Quantity of Money, Bonds, Commodity Inventories, and Capital." American Economic Review, Vol. 61, 335-349.

Finnery, J. D. (1980). "Real Money Balances and the Production Function." Journal of Money, Credit and Banking, Vol. 12, 666-671.

Fisher, I. (1922). The Purchasing Power of Money. New York: Macmillan and Company, Ltd.

Fisher, R. A. (1958). Statistical Methods for Research Workers. Thirteenth Edition. New York: Hafner Publishing Company, Inc.

Fisher, S. (1974). "Money and the Production Function." Economic Inquiry, Vol. 12, 517-533.

Fisher, W. D. (1969). Clustering and Aggregation in Economics. Second Edition. Baltimore: The Johns Hopkins Press.

Fraizer, W. (1964). "The Financial Structure of Manufacturing Corporations and the Demand for Money: Some Empirical Findings." Journal of Political Economy, Vol. 72, 176-183.

Friedman, M. (1956). "The Quantity Theory of Money--A Restatement." In M. Friedman, Ed., Studies in the Quantity Theory of Money. Chicago: University of Chicago Press, 6-22.

_________________ (1959). "The Demand for Money: Some Theoretical and Empirical Results." Journal of Political Economy, Vol. 67, 301-342.

_________________ (1970). "A Theoretical Framework for Monetary Analysis." Journal of Political Economy, Vol. 78, 190-212.

Frost, P. (1970). "Banking Services, Minimum Cash Balances, and the Firm's Demand for Money." Journal of Finance, Vol. 25, 1029-1039.

Fuss, M., D. McFadden, and Y. Mundlak (1978). "A Survey of Functional Forms in the Economic Analysis of Production." In M. Fuss and D. McFadden, Eds., Production Economics: A Dual Approach. Vol. I. Amsterdam: North Holland Publishing Company, 219-268.

Gabor, A., and I. Pierce (1958). "The Place of Money Capital in the Theory of Production." Quarterly Journal of Economics, Vol. 72, 537-557.

Goldberger, A. S. (1964). Econometric Theory. Second Edition. New York: John Wiley and Sons, Inc.

Goldfeld, S. (1976). "The Case of the Missing Money." Brookings Papers on Economic Activity, Vol. 3, 683-730.

Goldstein, M., and A. F. Smith (1974). "Ridge-type Estimators for Regression Analysis." Journal of the Royal Statistical Society, Series B, Vol. 36, 284-291.

Granger, C. W., and P. Newbold (1973). "Some Comments on the Evaluation of Economic Forecasts." Applied Economics, Vol. 5, 35-47.

_________________ (1974). "Spurious Regressions in Econometrics." Journal of Econometrics, Vol. 2, 111-120.

_________________ (1977). Forecasting Economic Time Series. New York: Academic Press.

Green, H. A. (1964). Aggregation in Economic Analysis. Princeton, New Jersey: Princeton University Press.

Grossman, H., and A. Policano (1975). "Money Balances, Commodity Inventories, and Inflationary Expectations." Journal of Political Economy, Vol. 83, 1093-1112.

Gunst, R. F., and R. L. Mason (1977). "Biased Estimation in Regression: An Evaluation Using Mean Square Error." Journal of the American Statistical Association, Vol. 72, 616-628.

Guttenberg, J., and R. Davis (1962). "Are Compensating Balance Requirements Irrational?" Journal of Finance, Vol. 17, 115-130.

__________ (1963). "Balance Requirements and Deposit Compensation." Journal of Political Economy, Vol. 71, 581-585.

Hall, R. E., and D. W. Jorgenson (1967). "Tax Policy and Investment Behavior." American Economic Review, Vol. 57, 391-414.

Hatanka, M. (1974). "A Simple Suggestion to Improve the Mincer-Zarnowitz Criterion for the Evaluation of Forecasts." Annals of Economic and Social Measurements, Vol. 3, 521-524.

Hoerl, A. E., and R. W. Kennard (1970a). "Ridge Regression. Applications to Nonorthogonal Problems." Technometrics, Vol. 12, 69-81.

__________ (1970b). "Ridge Regression. Biased Estimation for Nonorthogonal Problems." Technometrics, Vol. 12, 55-67.

Humphrey, D. B., and J. Moroney (1975). "Substitution Among Capital, Labor, and Natural Resource Products in American Manufacturing." Journal of Political Economy, Vol. 83, 57-82.

Intriligator, M. D. (1978). Econometric Models, Techniques, and Applications. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Johnson, H. (1962). "Monetary Theory and Policy." American Economic Review, Vol. 52, 335-384.

Johnson, H. (1969). "Inside Money, Outside Money, Income, Wealth, and Welfare in Monetary Theory." Journal of Money, Credit, and Banking, Vol. 1, 30-45.

Kakwani, N. C. (1967). "The Unbiasedness of Zellner's Seemingly Unrelated Regression Equations Estimators." Journal of the American Statistical Association, Vol. 62, 141-142.

Karni, E. (1973). "The Transactions Demand for Cash: Incorporation of the Value of Time Into the Inventory Approach." Journal of Political Economy, Vol. 83, 1216-1225.

Katzner, D. W. (1970). Static Demand Theory. New York: Macmillan Company.

Khan, M. S., and P. J. Kouri (1975). "Real Money Balances: An Omitted Variable from the Production Function." Review of Economics and Statistics, Vol. 57, 244-246.

Kmenta, J. (1971). Elements of Econometrics. New York: MacMillan Company.

Lange, O. (1936). "The Place of Interest in the Theory of Production." Review of Economic Studies, Vol. 3, 159-192.

Lau, L. J. (1974). "Comments on Applications of Duality Theory." In M. D. Intriligator and D. A. Kandrick, Eds., Frontiers of Quantitative Economics. Vol. II. Amsterdam: North-Holland Publishing Company, 176-199.

__________ (1978). "Testing and Imposing Monotonicity, Convexity, and Quasi-Convexity Constraints." In M. Fuss and D. McFadden, Eds., Production Economics: A Dual Approach. Vol. I. Amsterdam: North-Holland Publishing Company, 409-452.

Lehmann, E. L. (1959). Testing Statistical Hypothesis. New York: John Wiley and Sons, Inc.

Levhari, D., and D. Patinkin (1968). "The Role of Money in a Simple Growth Model." American Economic Review, Vol. 58, 713-758.

McCall, J. (1960). "Differences Between the Personal Demand for Money and the Business Demand for Money." Journal of Political Economy, Vol. 68, 353-368.

McFadden, D. (1978). "Estimation Techniques for the Elasticity of Substitution and Other Production Parameters." In M. Fuss and D. McFadden, Eds., Production Economics: A Dual Approach. Vol. I. Amsterdam: North-Holland Publishing Company, 73-118.

Madansky, A. (1964). "On the Efficiency of Three-Stage Least-Squares Estimation." Econometrica, Vol. 32, 51-56.

Maddala, G. S. (1971). "The Use of Variance Components Models in Pooling Cross Section and Time Series Data." Econometrica, Vol. 39, 341-358.

Malinvaud, E. (1970). Statistical Methods of Econometrics. Second Edition. Amsterdam: North-Holland Publishing Company.

Meltzer, A. H. (1963a). "The Demand for Money: A Cross-Section Study of Business Firms." Quarterly Journal of Economics, Vol. 77, 405-422.

__________ (1963b). "The Demand for Money: The Evidence from the Time Series." Journal of Political Economy, Vol. 71, 219-246.

Meltzer, A. H., and K. Brunner (1967). "Economies of Scale in Cash Balances Reconsidered." Quarterly Journal of Economics, Vol. 81, 422-436.

Miller, M. H., and D. Orr (1966). "A Model of the Demand for Money by Firms." Quarterly Journal of Economics, Vol. 80, 413-435.

__________ (1967). "An Application of Control Limit Models to the Management of Corporate Cash Balances." In J. Robichek, Ed., Financial Research and Management Decisions. New York: John Wiley and Sons, Inc., 169-185.

__________________ (1968). "The Demand for Money by Firms: Extensions of Analytic Results." Journal of Finance, Vol. 23, 735-759.

Mincer, J., and V. Zarnowitz (1969). "The Evaluation of Economic Forecasts." In J. Mincer, Ed., Economic Forecasts and Expectations. New York: National Bureau of Economic Reserarch, 124-127.

Moroney, J. (1972). "The Current State of Money and Production Theory." American Economic Review, Vol. 62, 335-343.

Nadiri, M. (1969). "The Determinates of Real Cash Balances in the U. S. Total Manufacturing Sector." Quarterly Journal of Economics, Vol. 83, 173-196.

Nelson, C. R. (1972). "The Prediction Performance of the F.R.B.-M.I.T.-Penn. Model of the U. S. Economy." American Economic Review, Vol. 62, 902-917.

Newbold, P., and C. W. Granger (1974). "Experience With Forecasting Univariate Time Series and the Combination of Forecasts." Journal of the Royal Statistical Society, Vol. 37, 131-146.

Niccoli, A. (1975). "Real Money Balances and Production: A Note." Review of Economics and Statistics, Vol. 57, 241-245.

Office of the Presidency (1960-1981). Economic Report of the President Transmitted to Congress. Washington: U. S. Government Printing Office.

Orr, D. (1970). Cash Management and the Demand for Money. New York: John Praeger Publishers.

__________________ (1971). "A Note on the Uselessness of Transaction Demand Models." Journal of Finance, Vol. 47, 1565-1572.

Parkin, M., and F. Edgar (1971). "The Optimal Quantity of Money, Bonds, Commodity Inventories, and Capital." American Economic Review, Vol. 61, 330-342.

Patinkin, D. (1965). Money, Interest, and Prices. Second Edition. New York: Harper and Row Publishing.

Pindyck, R. S., and D. L. Rubinfeld (1976). Econometric Models and Economic Forecasts. First Edition. New York: McGraw-Hill Book Company.

Porter, R., and E. Mauskopf (1978). "Financial Innovation and the Monetary Aggregates." Brookings Papers on Economic Activity, Vol. 1, 213-219.

__________________ (1979). "Cash Management and the Recent Shift in the Demand for Demand Deposits." (Unpublished manuscript, Federal Reserve Board of New York, New York.)

Prais, Z. (1975). "Real Money Balances as a Variable in the Production Function." Review of Economics and Statistics, Vol. 57, 243-244.

Rao, P., and Z. Griliches (1969). "Small-Sample Properties of Several Two-Stage Regression Methods in the Context of Autocorrelated Errors." Journal of the American Statistical Association, Vol. 64, 253-272.

Santomero, A. (1974). "A Model of the Demand for Money by Households." Journal of Finance, Vol. 29, 89-102.

Sargan, J. D. (1964). "Three-stage Least-Squares and Full Maximum Likelihood Estimates." Econometrica, Vol. 32, 77-81.

Saving, T. (1972). "Transactions Costs and the Demand for Money." Journal of Money, Credit, and Banking, Vol. 4, 245-259.

Seldon, R. (1961). "The Postwar Rise in the Velocity of Money: A Sectoral Analysis." Journal of Finance, Vol. 17, 483-555.

Shephard, R. W. (1953). Cost and Production Functions. Princeton, New Jersey: Princeton University Press.

________________ (1970). Theory of Costs and Production Functions. Princeton, New Jersey: Princeton University Press.

Short, E. (1979). "A New Look at Real Money Balances as a Variable in the Production Function." Journal of Money, Credit, and Banking, Vol. 11, 326-339.

Sinai, A., and H. Stokes (1972). "Real Money Balances: An Omitted Variable From the Production Function?" Review of Economics and Statistics, Vol. 54, 290-296.

________________ (1975). "Real Money Balances: An Omitted Variable From the Production Function--a Reply." Review of Economics and Statistics, Vol. 57, 247-252.

Sprenkle, C. (1966). "Large Economic Units, Banks and the Transactions Demand for Money." Quarterly Journal of Economics, Vol. 80, 436-443.

________________ (1969). "The Uselessness of Transactions Demand Models." Journal of Finance, Vol. 45, 835-847.

Standard and Poor's Corporation (1960-1981). Annual Dividend Records. New York: Poor's Publishing Company.

________________ (1960-1981). Corporation's Records. New York: Poor's Publishing Company.

________________ (1960-1981). Security Owner's Stock Guide. New York: Poor's Publishing Company.

Telser, L. G. (1964). "Iterative Estimation of A Set of Linear Regression Equations." Journal of the American Statistical Association, Vol. 59, 845-862.

Theil, H. (1954). Linear Aggregation of Economic Relations. Amsterdam: North-Holland Publishing Company.

____________ (1961). Economic Forecasts and Policy. Second Edition. Amsterdam: North-Holland Publishing Company.

____________ (1965). "The Analysis of Disturbances in Regression Analysis." Journal of the American Statistical Association, Vol. 60, 1067-1079.

____________ (1966). Applied Economic Forecasting. Chicago: Rand McNally and Company. Amsterdam: North-Holland Publishing Company.

____________ (1971). Principles of Econometrics. New York: John Wiley and Sons, Inc. Amsterdam: North-Holland Publishing Company.

Tinbergen, J. (1939). Econometrics. London: George Allen and Unwin, Ltd.

Tinsley, P. A., and B. Garrett (1978). "The Measurement of Money Demand." (Unpublished manuscript, Special Studies, Federal Reserve board of New York, New York.)

Tobin, J. (1956). "The Interest Elasticity of Transactions Demand for Cash." Review of Economics and Statistics, Vol. 38, 241-247.

Tsaing, S. (1969). "The Precautionary Demand for Money: An Inventory Theoretical Analysis." Journal of Political Economy, Vol. 77, 99-117.

U. S. Department of Labor, Bureau of Labor Statistics (1960-1981). Employment and Earnings. Library of Congress Catalog Number 70-11379. Washington: U. S. Government Printing Office.

____________ (1960-1981). Producer Prices and Prices Indexes. Library of Congress Catalog Number L53-140. Washington: U. S. Government Printing Office.

Uzawa, H. (1964). "Duality Principles in the Theory of Cost and Production." International Economic Review, Vol. 5, 216-220.

The Value Line Investment Survey. New York: Arnold Bernard and Company, Inc., 1960-1981.

Vogel, R., and G. Maddala (1967). "Cross Section Estimates of Liquid Asset Demand by Manufacturing Corporations." Journal of Finance, Vol. 22, 557-575.

Wallace, T. D., and A. Hussain (1969). "The Use of Error Components Models in Combining Cross Section and Time Series Data." Econometrica, Vol. 37, 55-72.

Wallis, K. (969). "Some Recent Developments in Applied Econometrics: Dynamic Models and Simultaneous Equation Systems." Journal of Economic Literature, Vol. 7, 771-796.

Weitzman, M. (1968). "A Model of the Demand for Money by Firms: Comment." Quarterly Journal of Economics, Vol. 82, 161-164.

Wenninger, J., and C. Sivesind (1979). "Changing the M1 Definition: An Empirical Investigation." (Unpublished manuscript, Federal Reserve Bank of New York, New York.)

Whalen, E. (1965). "A Cross-Section Study of Business Demand For Cash." Journal of Finance, Vol. 20, 425-439.

________________ (1966). "A Rationalization of the Precautionary Demand for Cash." Quarterly Journal of Economics, Vol. 80, 314-324.

Whiten, T. M. (1957). The Theory of Inventory Management. Second Edition. Princeton, New Jersey: Princeton University Press.

Zellner, A. (1962). "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias." Journal of the American Statistical Association, Vol. 57, 348-368.

________________ (1963). "Estimators for Seemingly Unrelated Regression Equations: Some Exact Finite Sample Results." Journal of the American Statistical Association, Vol. 58, 977-992.

Zellner, A., and H. Theil (1962). "Three-Square Least Squares: Simultaneous Estimation of Simultaneous Equations." Econometrica, Vol. 30, 54-78.

APPENDIXES

APPENDIX A

OPPORTUNITY COST OF MONEY

Overview. The text suggests using the return on the asset most like money to measure money's price. The production model of Saving (1972) suggests a perspective on this selection based on an investment's discount rate.

Opportunity Cost of Money. Consider a schedule of assets containing no risk. Only proportional transfer costs (C) of moving into and out of these securities exist. The one period return on a zero transfer cost asset equals r_0. From this base, a spectrum of returns for securities will exist, each one corresponding to a different level of transfer costs. In particular, an investor must receive $I_j + C_j I_j$ to compensate for the transfer costs of investing an amount I_j in the j^{th} asset. Since r_0 can be earned on an asset with no such costs, the rate of return on $j(r_j)$ must satisfy,

$$I_j(1 + r_j) = (I_j + C_j I_j)\ (1 + r_0) \tag{3-1}$$

which states that the total payment from I_j invested at r_j must equal the total payment from $I_j(1 + C_j)$ invested at r_0. The relationship between r_j and r_0 follows by simplification,

$$1 + r_j = (1 + C_j)\ (1 + r_0)$$

$$r_j = r_0 + C_j + C_j r_0$$

$$r_j = r_0 + C_j(1 + r_0)$$

and states that the return on j must equal the return on a zero transfer cost asset plus the future value of the transfer costs per unit. Such a formulation implies a different rate of return for every asset due to different transfer costs. An investor must select the proper rate to discount his investment from this spectrum.

An example limited to two rates can illustrate the proper choice. When investing in j, let r_k represent the largest foregone earnings and r_0 the smallest. The net present value (NPV) of asset j equals,

$$NPV_j = \frac{(1 + r_j)I_j}{1 + r_k} - (1 + C_j)I_j$$

if r_k represents the correct choice. Yet, this equation reduces to,

$$NPV_j = \frac{-(1 + C_j)\ I_j\ (1 + r_k) + (1 + r_j)\ I_j}{(1 + r_k)}$$

and, from Equation (3-1),

$$I_j(1 + r_j) = (1 + C_j)\ I_j\ (1 + r_0).$$

Thus,

$$NPV_j = \frac{-(1 + C_j)\ I_j\ (1 + r_k) + (1 + C_j)\ I_j\ (1 + r_0)}{(1 + r_k)}$$

$$NPV_j = \frac{-(1 + C_j)\ I_j - (1 + C_j)\ I_j r_k + (1 + C_j)\ I_j + (1 + C_j)\ I_j r_0}{(1 + r_k)}$$

$$NPV_j = \frac{(1 + C_j)\ I_j\ (r_0 - r_k)}{(1 + r_k)}$$

which exceeds zero if $r_0 > r_k$ and is negative when $r_k > r_0$. This contradiction arises because the total dollar earnings on all assets must be equal so that all net returns equal r_0. This result implies,

$$NPV_j = \frac{(1 + r_j)I_j}{(1 + r_0)} - (1 + C_j)I_j$$

$$NPV_j = \frac{(1 + C_j)I_j\ (1 + r_0)}{(1 + r_0)} - (1 + C_j)\ I_j$$

$$NPV_j = 0\ .$$

If NPV exceeded zero, the abnormally high earnings would induce entry driving net returns down to r_0. As a consequence, the lowest return asset of the same risk and maturity represents the opportunity cost of an investment. To discount cash flows from earnings on j with a rate of return which reflects those same costs, constitutes double counting. The different rates available on assets of identical risk and maturity reflect the different costs of transferring them into money. Thus, the opportunity cost of investing in money does not equal the high rate on a security containing premiums for substantial transfer costs. It equals the return on a zero transfer cost asset of the same risk and maturity. In fact, regardless of an asset's actual transfer costs, its opportunity cost equals the return on a zero transfer cost asset having the same risk and maturity.

In the more general case where fixed transfer costs (a) exist along with C, the determination of rates of return is more complex. Also, investors now prefer certain assets. In the previous case, all assets gave the same net return; thus, no advantage existed in buying large amounts of one asset. Now a preference arises due to the option of spreading fixed costs over a large asset purchase. Thus, the exact relationship between r_j and r_0 depends on the distribution of wealth among investors since large investors desire to purchase assets with large fixed and small proportional costs. The only certain relation between r_j and r_0 is that r_j must exceed r_0.

For an investment in asset j, the NPV now equals

$$NPV_j = \frac{(1 + r_j)I_j}{(1 + r_0)} - a - (1 + C_j)I_j$$

where NPV_j can exceed zero since $r_j > r_0$ by some unknown amount. This implies that a net return--net of all transfer costs--greater than r_0 can be earned on an asset. In the previous case, arbitrage caused all net returns to equal r_0 and, thus, NPV_j took a value of zero for all j.

This result no longer holds because of the disparity of wealth among investors. Those with a large amount of funds will prefer to purchase a large amount of an asset having high fixed and low proportional transfer costs. This action allows a big investor to spread fixed transfer costs which lowers transfer costs per dollar of investment. Yet, investors with small funding will select assets having smaller fixed costs to keep average transfer costs low. Again, a hierarchy of assets--in terms of returns--will emerge where each rate is determined by the marginal investor.

For any asset j, the last investor will purchase the smallest dollar amount of j and receive a net return just equal to the return on a zero transfer cost asset (r_0). All other investors in j purchase larger amounts which lowers their average transfer costs allowing them to earn a net return in excess of r_0. Traders, seeking positive net returns, develop a k^{th} asset in the hierarchy below j when a gap exists in the investment funds of the smallest investor in j and the largest investor in k. Such gaps create different assets and result in the hierarchy. For example, treasury bills have large fixed costs and small proportional costs whereas a commercial bank time deposit entails small fixed costs and no proportional costs. For the assets to coexist, the rate on treasury bills must exceed that on time deposits to compensate the marginal investor in treasury bills for the higher transfer costs. Thus, the presence of separate markets reflects the difference in average transfer costs to the marginal--the smallest--investor in each market. Yet, the net return--of all transfer costs--to the last investor in each market will equal r_0. Arbitrage will not eliminate the higher net returns earned by large investors in treasury bills since they possess a scarce resource--above average wealth. This analysis does not imply that the opportunity cost of money equals the rate of return on an asset actually held.

The opportunity cost of money equals the return on a transfer costless asset since the net return at the margin for all assets equals r_0. For example, if a savings account is deemed a transfer costless asset, then money's opportunity cost, for all investors, equals the return on this asset. To conclude that the true cost is the return made by the average investor, not the marginal one, abandons the concept of marginalism which forms the basis of market pricing.

APPENDIX B

EXTENSIONS OF THE INVENTORY MODEL

Overview. The inventory variants of Baumol (1952) and Miller and Orr (1966) provide testable specifications describing a firm's transactions balances. As Sprenkle (1966) argues, these analysis do not incorporate compensating balances. The first theoretical discussion below vindicates inventory models from his criticism and forms a general equilibrium framework describing the relationship between firms and banks. The second theoretical topic employs this framework to incorporate cash management techniques into inventory models.

Compensating Balances. These inventory theories explain only the amount of money which firms hold to conduct transactions. They do not account for cash balances arising to serve other purposes. The existence of such additional balances in large amounts would make inventory theories nearly useless as models of behavior.

One source, traditionally cited, of alternative money holdings is compensating balances. Banks require firms to hold idle demand deposits in compensation for the array of services they provide which include check processing, loan arrangements, financial advice, etc. Firms willingly enter into such arrangements to receive these beneficial services. In turn, compensating balances reward banks for the provision of banking services in two ways. First, the arrangement guarantees the bank a minimum demand deposit level. Since firms cannot hold less than this amount, banks can lend the full amount of compensating balances without risk of withdrawal. Second, the arrangement lowers the variance of total demand deposits allowing the bank to safely loan out a greater percent of total deposits. Since mutual gain exists, a market develops for these exchanges.

Miller and Orr (1966) incorporate this traditional view of compensating balances by having their model describe firm's money demand above some minimum level. The latter being determined outside the model through the interaction of firms and banks for banking services. For simplicity, Miller and Orr set this minimum equal to zero when developing their model. Yet, compensating balances in excess of zero will cause inventory models to miscalculate actual behavior. Sprenkle (1969) cites this reason for the poor predictive ability of the Baumol equation. Although compensating balances could account for such results, the traditional analysis contains serious theoretical pitfalls.

For one, firms could pay for banking services directly instead of using idle balances. In fact, because of reserve requirements both firms and banks would benefit if payment was made directly instead of indirectly. The Federal Reserve and/or prudent banking dictates that banks hold part of demand deposits in non-interest bearing form. Since no such restrictions apply to business, they can invest the entire compensating balance. Their earnings would more than match the implicit payment allowing both groups to receive a greater return than with the compensating balance arrangement. Thus, only an irrational businessman would hold involuntary balances as payment to banks. Furthermore, banks should show no preference to indirect as opposed to direct payment. In fact, they allow individuals the option of paying directly (per check) or indirectly (minimum balance). No reason exists not to offer such an option to firms. Thus, compensating balances are not <u>involuntary</u> holdings used as indirect payment to banks for services.

This conclusion does not imply that firms make no indirect payments for banking services. Firms could forego interest earnings on <u>voluntarily</u> held demand deposits used for transactions purposes in order to compensate banks. Banks gain the earnings and pay firms indirectly in the form of services. Such an arrangement would be the natural response to government price ceilings on demand deposits.

When hiring demand deposits from firms to make their "product", banks incur the cost of providing transactions services, which include making payments for firms, clearing checks, keeping records, etc. These costs are then netted out from interest earned when lending the deposits. If net returns exceed zero, banks can pay firms explicitly for employing their money. The payment must equal the alternative cost of the resource, otherwise the firm will use its money in the alternative. In the absense of price ceiling regulation, the payment could take explicit form. With price controls, the payment must still be made, but now it must have an implicit form. Banking services constitute black market payments (in addition to absorbing transactions costs) needed to acquire a factor of production. This analysis implies that compensating balances are a portion of voluntarily held transactions balances and, thus, do not invalidate inventory theories. In other words, the bank's "requirement" of a minimum balance does not constitute a binding constraint on firms. Actually, banks compensate firms in the analysis. This arrangement fits naturally with a general equilibrium framework developed below.

A more serious theoretical pitfall of the traditional view of compensating balances is its inability to deal with the fundamental question of why firms use bank money as opposed to government money. The answer lies in a framework inconsistent with the traditional view of compensating balances.

Consider a business sector where all firms manage cash using an inventory approach and, thus, the cash balances of every firm exhibit economies of scale. This property implies certain relationships between firms whose descriptions requires a general equilibrium framework. As a firm acquires larger transactions levels, the amount of money required to conduct a dollar of transactions declines. In other words, economies of scale imply that the average cost of using money for transacting (ATRC) falls as size increases. Businesses should take advantage of such savings either by merger or internal growth. Alternatively, they could centralize their balances in one firm who could then specialize in receipt-payment operations. The bank, by combining these balances, could bring substantial savings in transactions cost to each firm.

The simple Baumol model can illustrate the extent of such savings from either growth or consolidation. The total cost of using cash for transactions purposes - total transactions costs - equal,

$$TTRC = b\left(\frac{T}{C}\right) + r\left(\frac{C}{2}\right)$$

where the first right hand term measures the transfer cost between interest bearing assets and cash. The second right hand expression represents the foregone interest earnings when using cash. Since average money holdings in this model equal,

$$M = \frac{C}{2} = \left(\frac{bT}{2r}\right)^{1/2}$$

substitution gives,

$$TTRC = bT\left(\frac{r}{2bT}\right)^{1/2} + r\left(\frac{bT}{2r}\right)^{1/2}$$

which reduces to,

$$TTRC = \left(\frac{b^2T^2r}{2bT}\right)^{1/2} + \left(\frac{bTr^2}{2r}\right)^{1/2}$$

$$TTRC = \left(\frac{bTr}{2}\right)^{1/2} + \left(\frac{bTr}{2}\right)^{1/2}$$

$$TTRC = \left(\frac{bTr}{2}\right)^{1/2} (1 + 1)$$

$$TTRC = 2 \left(\frac{bTr}{2}\right)^{1/2}$$

$$TTRC = (2bTr)^{1/2}.$$

Transactions cost per unit of transactions - average transactions costs equal,

$$ATRC = \frac{(2bTr)^{1/2}}{T}$$

$$ATRC = \left(\frac{2br}{T}\right)^{1/2}$$

which declines monotonically as T increases.

With respect to consolidation, if n firms of equal size exist, the total transactions cost for each appears as,

$$TTRC_n = (2bTr)^{1/2}$$

implying that the total transactions cost for the group equals,

$$TTRC = n(TTRC_n) = n(2bTr)^{1/2}.$$

However, if a bank assumes the balances of all n firms, then its average balance equals,

$$M = \frac{c}{2} = (\frac{bnT}{2r})^{1/2}$$

and total transactions costs become,

$$TTRC = n^{1/2} (2bTr)^{1/2}.$$

Total transactions cost per firm would equal,

$$TTRC_n = \frac{TTRC}{n}$$

$$TTRC_n = (\frac{2bTr}{n})^{1/2}$$

which is less than TTRC in the non-consolidation case and declines as n increases.

Firms can acquire the cost saving of economies of scale either by internal consolidation--growth--or external consolidation--using bank money. Other things equals, firms should grow (through merger or internally) to the point where the decline in average transactions cost and the increase in average production cost (APC) due to a marginal growth are equal. The increase in APC, from expansion, occurs due to diseconomies of size. In Figure 1, long run average production costs fall until an output level of Q_o. Expansion past that point by internal growth or merger brings on management difficulties which more than offset any economies due to greater specialization. Yet, the decline in ATRC, as shown in Figure 2, continues monotonically as output and, hence, transactions levels continue to increase. Total average cost (AC), depicted in Figure 3, in the long run equals the sum of long run APC and ATRC. It declines through an output level of Q_o since both APC and ATC decline.

From Q_0 to Q_1, AC continues to fall since ATRC falls faster than APC rises. Past the output level of Q_1, AC begins to increase as the diseconomies of size overcome any further transactions economies of size, implying a limit to the motive for expansion. Finally, this analysis is consistent with both conglomerate mergers and any empirical evidence of plants in excess of minimum efficient production size.

For firms to gain these cost savings through external consolidation, an institution must exist which stands willing and able to sell transactions services to firms. Banks represent such specialized institutions. Their willingness and ability to provide transactions services stems from the profit which arises if banks either have lower average transactions costs than firms or can earn higher returns on investment than firms. These possibilities will be explored in turn.

As argued previously, in order to retain demand deposits banks must pay an amount equal to firm's opportunity cost of money which equals the return on money's best alternative. For all but the marginal firm, this return is less than average transactions cost since the latter includes transfer costs as well as foregone interest. If this condition did not exist, firms would not hold money balances. Now, a bank's ATRC are less than any single firm with deposits in the bank due to economies of size. Thus, a bank earns a residual on demand deposits even if they only earn a return on investments equal to the firm's opportunity cost of money. Actually, banks not only have lower costs than firms, they also earn greater returns than firms on deposits.

Banks specialize in making high risk (high return) loans to users--individuals and small businesses--willing to pay premium rates. These rates compensate banks for their middleman role between demanders and

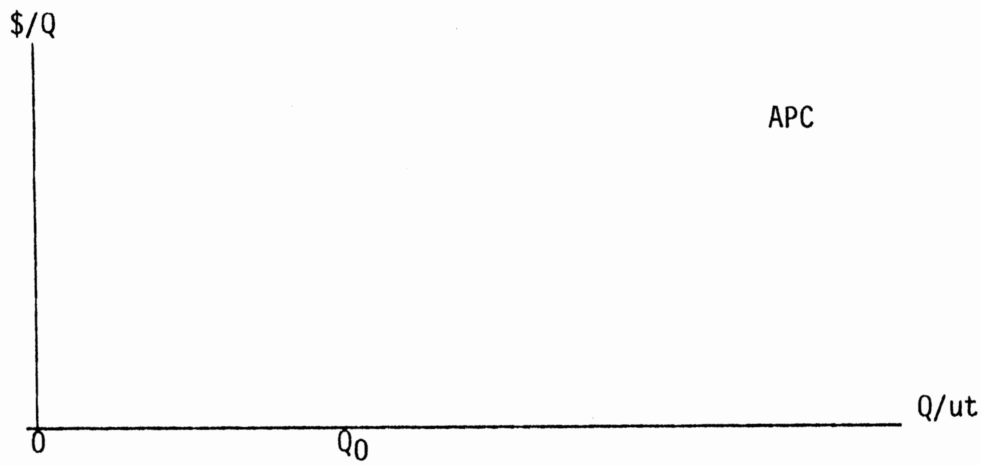

Figure 1. Average Production Costs

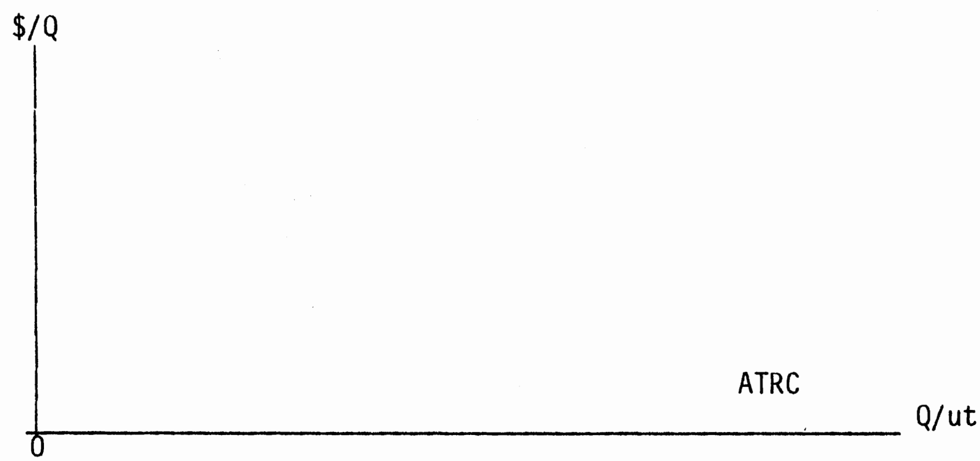

Figure 2. Average Transactions Costs

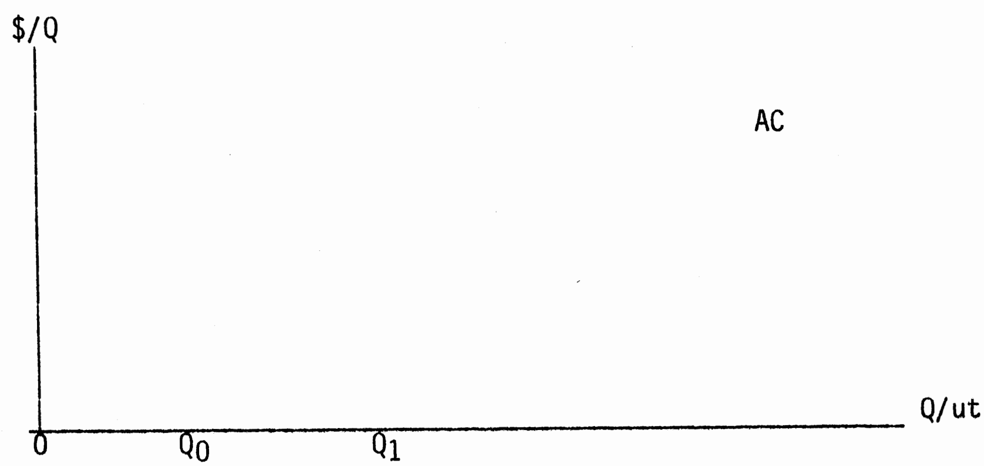

Figure 3. Average Costs

suppliers of loanable funds. Firms, in general, do not engage in such activity and, thus, accept lower returns on their securities. Furthermore, banks will have cash balances with relatively lower variances compared to an individual firm holding the same size average balance. The impact of this on earnings can be viewed within either the Baumol or Miller-Orr model.

In the Baumol model, a firm's cash balance equals zero at the start of a time period. During the period, an optimum number of equal amount withdrawals (C) from assets into cash occur. After cash holdings are replenished, the firm spends them at a constant rate until they equal zero. Then another withdrawal is made and the cycle repeats. This process results in a sawtooth cash balance whose average level equals one-half of the optimum cash withdrawal. Because many such transfers take place over a period, a firm cannot invest the full amount of average balances in earning form. The cash balance of a bank having the same average balance could vary between two extremes.

If all firms deposit and withdraw money simultaneously, a bank's cash balances will follow the same sawtooth process as the individual firm. However, if firms of equal size make evenly distributed deposits and withdrawals, banks cash balances will remain constant at a level equal to average balances. In this case, banks could lend the full amount of deposits at all times and, thus, increase earnings above the revenue of the individual firm. In fact, as long as some offsetting occurs, bank earnings will exceed those of an individual firm having the same average balance. The same qualitative conclusion results when using the Miller-Orr model.

In this case, firms have cash balances which vary randomly between an upper and a lower bound. This process (ε) has an expected value of zero--no systematic drift occurs--and a constant variance. Combining such a process with the asymmetric optimal control rules, the firm's average balances equal μ = ntm with a variance of σ^2 = ntm^2. A bank would consolidate the balances of two identical firms so that average balances are also μ = ntm. The process generating the bank's cash flows would be the sum of two random walks (ε_1 and ε_2). Each having zero mean, thus, their sum also has a zero mean. However, the variance would equal,

$$\mathrm{Var}(\varepsilon_1 + \varepsilon_2) = \mathrm{Var}(\varepsilon_1) + \mathrm{Var}(\varepsilon_2) + 2\ \mathrm{Cov}(\varepsilon_1\varepsilon_2)$$

$$\mathrm{Var}(\varepsilon_1 + \varepsilon_2) = 2\ \mathrm{Var}(\varepsilon_1) + 2\ \mathrm{Cov}(\varepsilon_1\varepsilon_2)$$

$$\mathrm{Var}(\varepsilon_1 + \varepsilon_2) = \mathrm{Var}(\varepsilon) + 2\ \mathrm{Cov}(\varepsilon_1\varepsilon_2)$$

where ε_i represents the random walk generating the two identical firm's cash balances. If the Cov($\varepsilon_1\varepsilon_2$) takes a negative value, the bank's variance will be less than the variance of an equal sized firm. This occurs since the two random walks would offset each other, allowing the bank to gain greater earnings through larger investments. Because banks have both lower costs per unit and higher earnings per unit than firms, they stand willing and able to offer transactions services to firms.

On the demand side of the market, firms desire to use some form of money to obtain transactions services. Two close substitutes exist. When using the fiat money of the government firms absorb all transactions costs. On the other hand, banks take on part of these costs if firms use demand demands. In addition to lowering a firm's ATRC, holding bank money entitles the firm to receive banking services. For these reasons, firms find demand deposits a better money and, thus, desire to hold large amounts relative to cash. The lowering of ATRC by consolidation and the increasing of interest earnings due to loan specialization provide funds for mutual gain to banks and firms resulting in the formation of a market.

In summary, the general equilibrium framework has the following implications. Banks arise as firms specializing in transactions services and loan middleman activity due to the potential for profit. Firms gain when keeping transactions balances in the form of bank money since it lowers ATRC and results in indirect payment via bank services. This payment by banks, compensating firms for providing loanable funds, is implicit due to government price ceilings on demand deposits. Also, given reserve requirement regulation, compensating balances must be part of voluntarily held transactions balances. They represent a nonbinding constraint since firms will hold demand deposit balances in excess of a compensating balance to obtain transaction cost savings and banking services. Thus, inventory models are not invalidated by the compensating balance arrangement. Finally, the framework can encompass the advances in cash management techniques.

<u>Cash Management Techniques Adjustments</u>. Recently the relationship between firms and banks has changed due to the expanded use of three groups of cash management techniques (CMT). Mutual gain has fostered a tremendous growth in the use of immediately available funds by both banks and firms, altering demand deposit relationships. Also, certain regulatory changes by the Federal Reserve have reduced firm's transfer costs and, thus, shifted their money demand functions. A final group of CMT--lock boxes, remote disbursements, etc.--provide ways to lower the variance of cash balances with a resulting shift in money demand. Each

group and its effect on inventory models will be outlined in turn.

The deterministic inventory variant has money demand dependent on transfer costs, transactions amounts, and the interest rate. The stochastic inventory variant postulates the same determinates except for replacing transactions amounts with daily cash flow variance. The emergence of each group of CMT affects a different independent variable in these specifications.

The creation of repurchase agreements (RP) as a secured immediately available fund (IF) provides a practical measure of money's opportunity cost. Because RP represent a riskless, near zero transfer cost asset, its return approximates what firms give up to use money. The only better alternative--interest bearing checkable deposits--have explicit returns subject to price ceilings, making them useless as a measure of opportunity cost. While the action of buyers and sellers set the return of RP, those rates are not readily available. Yet, a suitable proxy does exist.

RP yield should hold a proportional relationship to the federal funds rate (FFR), allowing the latter to provide a reasonably accurate proxy. Since the Federal Reserve prohibits non-bank traders in the federal funds market, only banks can participate in both markets. Their action will integrate the markets by establishing a cost based differential between the RP rate and the FFR. Any time trading alters this differential, banks can profit by arbitrage. Such counter buying or selling re-established the cost based wedge implying that the two rates of returns will move together. In addition to suggesting a proxy for money's opportunity cost, this group of CMT provides some implications which inventory models can test.

As argued by Tinsley and Garrett (1978), the introduction and growth of RP is not unique but fits a trend beginning in the 1960s. For every economic expansion since that time, interest rates on non-bank short term securities rise relative to bank deposit rates due to the price ceiling on bank deposit rates. The effect on money demand can be viewed within the Miller-Orr three-asset model.

This approach allows firms to hold a hierarchy of assets along with money. On the scale nearest cash lie shorts--money's closest alternative. As argued previously, checkable and savings deposits at banks fill this role in reality. Moving toward the spectrum's other end leads to longer maturity assets with increasing levels of both risk and transfer costs, called longs. Firms determine the optimum amount of money based on the return of and transfer costs into shorts. The preferred short balance depends on own rates, cash holdings, and the return on and transfer costs into longs. Finally, long holdings depend on own rates and rates for and transfer costs into shorts. Such a structure induces firms to shift wealth only between neighboring assets. For example, when firms desire to increase long holdings at the expense of other assets, the restructuring occurs in the following way.

The funds necessary to make long purchases come from short holdings not cash. As short balances decline, the relationship between shorts and money changes inducing firms to move cash into shorts. Thus, the model postulates that changes in long balances affect cash indirectly through alterations in returns and transfer costs of shorts. This framework explains the emergence of new bank liabilities as follows.

In an economic expansion, the rate on longs increases relative to the fixed return on bank deposits--shorts. Firms restructure by pulling funds out of shorts to purchase longs, then partially restocking short holdings at the expense of demand deposits. Banks cannot stop this exodus by making their deposits competitive with longs because of price ceilings. Instead, they respond by creating a new security which can recapture the lost funds. Such an asset must provide a unique property to accomplish the task. In other words, it must fill a gap in the hierarchy of assets.

In the late 1960s, Euro-dollars arose in response to firms' portfolio adjustments. Trading grew to a peak in 1968, coinciding with the peak on short term non-bank investments. During the expansion of the early 1970s, certificates of deposits and RP were introduced. Their growth peaked during 1973, corresponding to the peak in rates on non-bank investments. The introduction of these assets give firms a new alternative in the hierarchy which will permanently alter the relationship between money and shorts as firms begin to use the new asset. For a given amount of wealth, this restructuring means firms hold less money even though the variables of money demand have not changed. In econometric terms, introducing a new asset will cause structural change of the money demand function. The second group of CMT have also altered a parameter value of inventory models.

Certain regulatory changes have altered the cost of making transfers between assets. For one, corporations can now hold commercial bank savings and checkable accounts which have very low transfer costs compared to alternative assets. Also, the transfer cost of those deposits has been lowered by the acceptance of telephone and pre-authorized transfers between savings and demand deposits. Finally, the Fed now allows additional branch offices resulting in lower nonmonetary costs of transfers. As a result of these CMT firms desire to hold less money as a function of own price. This restructuring alters the relationship between banks and firms. The M-O three-asset model can outline the entire process and lead to a proxy for transfer costs.

As the transfer costs of bank deposits fell, firms desired to hold less money and more shorts. Since long balances depend on short holdings, this movement caused an imbalance between shorts and longs which firms corrected by purchasing additional longs. Banks, faced with the depletion of both demand and interest bearing deposits, introduced RP to re-attract the funds. A subsequent increase in trading of RP should occur which parallels the decline in transfer cost as banks make RP use increasingly attractive to channel the additional lost deposits back into banks in the form of RP. Thus, the inverse of RP trading volume can serve as a proxy for transfer costs over the sample period. This idea can also extend to the third group of CMT.

These techniques lower transfer costs, in general, not just on bank deposits. The decline in these costs for longs causes firms to react by shifting wealth into longs from shorts. This movement upsets the balance between shorts and money inducing firms to shift funds from demand deposits into shorts. As in the previous case, banks respond to the outflow by offering a new asset. Thus, a decline in transfer costs of either shorts or longs can be modeled within the Miller-Orr three-asset model. The resulting analysis justifies using the inverse volume of RP trading as a proxy for transfer costs. The third group of CMT can be modeled differently by using the original Miller-Orr approach.

Lock boxes, remote disbursements, electronic transfers, consolidated balances, etc., allow firms to control the variance of cash flows. As argued by Porter and Mauskopf (1979), prior to the emergence of these techniques firms treated the variance as a given parameter. Now, they give firms another control variable in managing cash balances, thus, money holdings depend on the optimum amount of CMT purchased. In formal terms, this amount is embodied in the addition of perceived variance to the simple Miller-Orr model.

Purchasing CMT allows firms to lower perceived as well as actual cash flow variances. With respect to the former, lock boxes affect cash receipts by decreasing the lag between sales and collected balances. Remote disbursements increase the lag between purchases and debiting of the firm's account. Altering the pattern of receipts and disbursements lower perceived variance because they decrease the time that a receipt or disbursement can show a nonzero value. Also these techniques may lower optimum demand deposits further as firm's float increases. This addition allows firms to safely invest more demand deposits in interest earning form.

Firms can lower actual variances through a number of techniques. Wire transfers and deposits transfer checks make concentration accounts feasible. These accounts consolidate receipts and disbursements from many deposits into one large deposit. The variance of two small deposits, in the Miller-Orr model, would equal $(\sigma^2)^{1/3} + (\sigma_2^2)^{1/3}$ while a consolidated account would have a smaller variance of $(\sigma_1^2 + \sigma_2^2)^{1/3}$. Thus, firms hold less money even in this extreme case where pooling does not lead to an offsetting receipt-payment pattern. Also, many CMT provide better information about cash flows inducing firms to invest a greater amount of demand deposits.

Automated retrieval systems allow firms to obtain information quickly about remote deposits, which promotes more rapid consolidation and investment. Payable through drafts remove uncertainty about the timing of check clearing which releases funds formerly held in anticipation of clearing. Zero balance accounts allow payment after a check comes in for clearance, freeing more funds for investment. Wire, telephone, and computers facilitate information retrieval on cash flows. Banks provide firms much information about average disbursement float and funds needed to cover checks clearing in any given day. Finally, forecasting methods on cash positions aid businesses in managing their portfolios. These devices result in more accurate decisions about managing cash as reflected in lower perceived variances.

In summary, the simple inventory models are significantly altered by each group of CMT. The opportunity cost of money and transfer costs both receive an accurate proxy. Also, the variance of cash flows has undergone significant change due to CMT. Methods to test the implications of this work appear in the following section.

Summary. Inventory theories postulate economies of scale in cash holdings. The resulting decline in average transactions costs induce firms to expand beyond minimum average production cost output. More importantly, economies provide profit incentive for banks to emerge since they can consolidate transactions balances of many firms. Government intervention affects this basic relationship in two ways. Price ceilings on demand deposits force banks to hire money by making in-kind payments. These give firms incentive to use bank as apposed to government money. Reserve requirements make it irrational for firms to hold involuntary compensating balances. They are part of the voluntarily held transactions balances and, thus, do not constitute a binding constraint on firms. Inventory models, vindicated from the compensating balance critique, can accurately incorporate the expanding relationship between firms and banks.

The emerging use of cash management techniques suggests the following additions to money demand. First, not only does money's price receive an accurate proxy, so does transfer costs. During economic expansion of the 1970s, the returns on non-bank investments increased while their transfer costs declined. Banks could not make traditional deposits competitive due to price controls. Instead, they introduced a new liability which filled a gap near money in the liquidity hierarchy. The volume of repurchase agreement trading increased as the returns of and transfer costs on alternative investments rose and fell, respectively. Thus, the inverse of repurchase agreement volume provides an accurate proxy for transfer costs during the time span under study. Second, the effect of cash management techniques can be modeled as lowering both perceived and actual cash flow variance. The Miller-Orr model can capture this change after a minor adjustment.

APPENDIX C

AGGREGATION ANALYSIS

Overview. This appendix presents a detailed outline of the general aggregation problem.

Aggregation Analysis. Economic theories of money demand must use empirical data to test their explanatory and predictive usefulness. Previous studies have employed aggregated data in this endeavor which requires postulating a macrorelationship. The reasons for using a macrorelation are simplicity and the lack of firm level data. Yet, this process cannot be generally valid since relations between aggregates result from many decisions of individuals. The logical step preceding the use of a macrorelation would be to test whether individual decisions, in some sense, form a stable aggregate relation.

Money demand functions describe individual firm behavior while macrorelations must be derived from these microfoundations. The translation tion from one level to the next must have an explicit form--an aggregation. This process is vital for both economic and statistical aspects of testing.

For this study, the problem takes a form limited to a single linear macrorelation derived from a set of linear microrelations by linear aggregation. If a macrorelation does not have a consistent form with respect to its micro counterparts, using the former as a true model constitutes a specification error. Only in the absence of this circumstance can a macrorelation properly test firms' money demand functions. Firm level data provides a basis to test for these conditions.

Employing the analysis of Theil (1954), consider a group of firms whose demand for cash depends on their level of transactions. For each of N firms, money demand is,

$$m_{ci} = \alpha_c + \beta_c t_{ci} + \varepsilon_{ci} \qquad c = 1 \ldots N$$

$$i = 1 \ldots n.$$

Unfortunately, only data on total money demand and total transactions for the group of firms exist. That is, macrovariables' measuring per firm amounts appear as,

$$M_i = \frac{1}{N}\sum_c m_{ci} \qquad\qquad T_i = \frac{1}{N}\sum_c t_{ci}$$

which describes the method of aggregation given, in this case, by the data. This procedure involves a theoretical contradiction since certain sums may not have the same relation to other sums even when the relation holds for all components. Yet, previous econometric studies proceed with the macrorelation as a linear sum of the microrelations. One facet of aggregation theory deals with the validity of this transformation from micro to macro levels.

Faced with a microrelation and only macro data, aggregation theory constructs a macrorelation in terms of per firm observations corresponding to the microrelation. Combining the macrovariables and aggregation method with the microrelation gives,

$$M_i = \frac{1}{N}\sum_c \alpha_c + \frac{1}{N}\sum_c \beta_c\ t_{ci} + \frac{1}{N}\sum_c \varepsilon_{ci}$$

as the correct macrorelation. It describes per firm money demand as a function of three terms. The first, a constant, equals the arithmetic mean of the intercepts of all individual firm demand equations. The last expression represents a random disturbance as the arithmetic mean of the microrelation's disturbances. The middle term is the weighted sum of firms transactions levels where the weights equal $\frac{\beta_1}{N} \ldots \frac{\beta_N}{N}$. Such an equation does not describe per firm demand in terms of per firm transactions unless $\beta_1 = \ldots = \beta_N = \beta$. In this case, M_i equals,

$$M_i = \frac{1}{N}\sum_c \alpha_c + \beta\ T_i + \frac{1}{N}\sum_c \varepsilon_{ci}$$

which is a linear equation in the macrovariables. This expression is called an incorrect macrorelation since, in general, it does not equal the correct macrorelation. In fact, if the coefficient parameters differ for individual firms, specification error occurs when using aggregate data since it implies estimating with the incorrect macrorelation. The consequences of such an error can be seen by analyzing the least square estimate of β as,

$$b = \frac{\sum_i (T_i - \overline{T})\ (M_i - \overline{M})}{\sum_i\ (T_i - \overline{T})^2}$$

where $\overline{T}$ and $\overline{M}$ represent the arithmetic averages, over n observations, of the corresponding macrovariables. To evaluate b, an expression for $M_i - \overline{M}$ must be obtained. $\overline{M}$ comes from carrying out the summation over n for the correct macrorelation. Thus,

$$\overline{M} = \frac{1}{N}\sum_c \alpha_c + \frac{1}{N}\sum_c \beta_c\ \overline{t}_c + \frac{1}{N}\sum_c \overline{\varepsilon}_c.$$

Subtraction of $\overline{M}$ from M_i gives

$$M_i - \overline{M} = \frac{1}{N}\sum_c \beta_c(t_{ci} - \overline{t}_c) + \frac{1}{N}\sum_c\ (\varepsilon_{ci} - \overline{\varepsilon}_c)$$

then by substitution, b equals,

$$b = \frac{\sum_i (T_i - \overline{T})\sum_c \beta_c(t_{ci} - \overline{t}_c)}{N\sum_i (T_i - \overline{T})^2} + \frac{\sum_i (T_i - \overline{T})\sum_c\ (\varepsilon_{ci} - \overline{\varepsilon}_c)}{N\sum_i (T_i - \overline{T})^2}\ .$$

Assuming nonstochastic values for transactions, the second righthand term's expectation vanishes since the mean of all disturbance terms equals zero. Thus, the expected value of b takes the form,

$$E(b) = \sum_c P_c \beta_c$$

where

$$P_c = \frac{\sum_i (T_i - \bar{T})(t_{ci} - \bar{t}_c)}{N \sum_i (T_i - \bar{T})^2}.$$

P_c represents the slope estimate of the auxiliary regression equal to,

$$\frac{1}{N} t_{ci} = a_c + \theta_c T_i + U_c$$

where a_c is a constant and U_c a disturbance term. This regression has the explanatory variable from the correct macrorelation as a dependent variable and the explanatory variable from the incorrect macrorelation as an independent variable. θ_c will differ from $1/N$ when the microparameters are not all identical. Thus, the slope of the incorrect macrorelation, which actually undergoes estimation has an expectation equal to a weighted sum of the microparameters--$\beta_1 \ldots \beta_N$--with weights of $P_1 \ldots P_N$. Finally, these weights sum to unity as,

$$\sum_c P_c = \frac{\sum_c [\sum_i (T_i - \bar{T})(t_{ci} - \bar{t}_c)]}{N \sum_i (T_i - \bar{T})^2}$$

and bringing the summation operator inside,

$$\sum_c P_c = \frac{\sum_i [(T_i - \bar{T}) \sum_c (t_{ci} - \bar{t}_c)]}{N \sum_i (T_i - \bar{T})^2}.$$

By definition, $T_i = \frac{1}{N} \sum_c t_{ci}$ and $\bar{T} = \frac{1}{N} \sum_c \bar{t}_c$.

Thus,

$$\sum_c P_c = \frac{\sum_i [(T_i - \bar{T})\ N\ (T_i - \bar{T})]}{N \sum_i (T_i - \bar{T})^2}$$

$$\sum_c P_c = 1 \ .$$

These results imply that the expectation of b will generally not equal β. For most firms, transactions levels change in the same direction with per firm transactions implying that t_{ci} and T_i are positively correlated. Thus, the auxiliary regression will result in a value for P_c greater than $^1/N$. Such a firm behaves in a manner which has a positive effect on the expectation of b. If t_{ci} and T_i show negative correlation, the firm's action affects b in an inverse manner. Yet, β should always exceed zero. In this case, the aggregation process introduces an aggregation bias into the macrorelation which manifests itself by allowing $E(b) \neq \beta$.

Following Green (1964), this bias can be analyzed using the correlation contained in the auxiliary regressions. Specifically, the covariance of a microparameter (β_c) with the macroparameter estimate from the auxiliary regression (P_c) appears as,

$$\mathrm{Cov}(\beta_c P_c) = \frac{1}{N} \sum_c [(\beta_c - \overline{\beta}_c)\ (P_c - \overline{P}_c)]$$

where

$$\overline{\beta}_c = \frac{1}{N} \sum_c \beta_c$$

$$\overline{P}_c = \frac{1}{N} \sum_c P_c = \frac{1}{N}.$$

Thus, by substitution,

$$\mathrm{Cov}(\beta_c P_c) = \frac{1}{N} \sum_c (\beta_c P_c - \overline{\beta}_c P_c - \frac{1}{N} \beta_c + \frac{1}{N} \overline{\beta}_c)$$

$$\mathrm{Cov}(\beta_c P_c) = \frac{1}{N} \sum_c [\beta_c P_c - \frac{1}{N} \beta_c] + \frac{1}{N} \sum_c [\frac{1}{N} \overline{\beta}_c - \overline{\beta}_c P_c]$$

$$\mathrm{Cov}(\beta_c P_c) = \frac{1}{N} (\sum_c \beta_c P_c) - \frac{1}{N} \frac{N}{N} \overline{\beta}_c + \frac{1}{N} [\frac{1}{N} \sum_c \overline{\beta}_c - \frac{1}{N} \sum_c \overline{\beta}_c]$$

$$\mathrm{Cov}(\beta_c P_c) = \frac{1}{N} (\sum_c \beta_c P_c) - \frac{1}{N} \overline{\beta}_c$$

Solving for the macroparameter's expectation gives,

$$E(b) = \sum_c P_c \beta_c = \overline{\beta}_c + N\ \mathrm{Cov}(\beta_c P_c)$$

where $\overline{\beta}_c$ represents an unbiased estimate, given the aggregation process used. The remaining term measures the bias introduced when that process causes misspecification. Consider firms with $\beta_c > \overline{\beta}_c$ whose transactions levels (t_{ci}) increase rapidly compared to the average (T_i); thus, $P_c >$

$^1/N$. Such a disproportionate relationship implies $Cov(\beta_c P_c) > 0$ resulting in $\beta > \overline{\beta}_c$. The macroparameter for money demand will exceed the level as given by the mean of separate microparameters. The bias term highlights the complications arising between micro and macro levels.

These problems give rise to contradictions when predicting with the microrelations or macrorelations. For example, if t_{ci} increases by Δt_{ci}, the resulting rise in aggregate transactions equals $\Delta T_i = \sum_c \Delta t_{ci}$. Yet, the effect on M_i can be traced from two paths. One uses the microrelations as,

$$\Delta_1 M_i = \sum_c \Delta m_{ci} = \sum_c \beta_c \Delta t_{ci} \ .$$

The other employs the macrorelation to achieve,

$$\Delta_2 M_i = \beta \Delta T_i = \beta \sum_c \Delta t_{ci}$$

where $\beta = E(b) = \sum_c \theta_c \beta_c$ and θ_c is the slope parameter of the auxiliary regression. In the general case, $\Delta_1 M_i$ does not equal $\Delta_2 M_i$. Two important aspects of this contradiction exist. First, even when all β_c remain constant (and likewise $\Delta_1 M_i$) different movements of transactions over time can change β and, hence, change the prediction of $\Delta_2 M_i$. For example, suppose each t_{ci} varies in a fixed proportion to T_i, although the proportion can be different for each firm. Each proportion equals the slope parameter (θ_c) of the respective auxiliary regressions as, $t_{ci} = \theta_c T_i + U_{ci}$ where the absence of the constant term reflects the fixed proportion assumption. Averaging over n observations,

$$\theta_c = \frac{\overline{t}_c}{\overline{T}}$$

which remains constant. Using the fact that $\beta = \sum_c \theta_c \beta_c$,

$$\beta = \frac{1}{\overline{T}} \sum_c \beta_c \overline{t}_c \ .$$

Thus, the estimation of β, after aggregation from β_c, gives greater weight to firms with larger transactions levels (t_c). This condition causes no problem so long as increases in transactions are proportional to the level of transactions for each firm. Formally,

$$\frac{\Delta t_{ci}}{\overline{t}_c} = \frac{\Delta T_i}{\overline{T}} \ .$$

Solving this expression for Δt_{ci} and substituting into $\Delta_1 M_i$ gives,

$$\Delta_1 M_i = \frac{\Delta T_i}{\overline{T}} \sum_c \beta_c \overline{t}_c = \beta \Delta T_i = \Delta_2 M_i$$

However, any time the distribution of transactions changes when aggregate transactions change, the auxiliary regressions become $t_{ci} = \alpha_c + \theta_c T_i + U_{ci}$. Thus,

$$\frac{t_{ci}}{T_i} = \theta_c + \frac{\alpha_c}{T_i}$$

which implies that transactions of large firms (large α_c) change by a greater proportion than that for small firms (small α_c) when aggregate transactions increase. That is, Δt_{ci} no longer stays fixed for each firm as aggregate transactions (T_i) change which causes $\Delta_1 M_i$ to differ from $\Delta_2 M_i$. This discrepancy increases when transactions for a firm move inversely with T_i since θ_c will take a negative value. In that case, the aggregate β can be larger or smaller than any of the individual β_c.

The second aspect of contradiction between $\Delta_1 M_i$ and $\Delta_2 M_i$ is that the distribution of Δt_{ci} to all firms can be disproportionate as opposed to the fixed proportion case above. The only restriction of Δt_{ci} states that $\sum_c \Delta t_{ci} = \Delta T_i$. Two different allocations of a given ΔT_i will cause $\Delta_1 M_i$ to change even when ΔT_i and $\Delta_2 M_i$ do not. Very stringent requirements are necessary to ensure no contradiction irrespective of the distribution of Δt_{ci}. Even when the $Cov(P_c \beta_c) = 0$, contradictions can arise due to nonuniform distribution.

The sufficient condition to avoid any contradiction between microrelations and macrorelations requires,

$$\beta_c = \frac{1}{N} \sum_c \beta_c = \overline{\beta}_c \text{ for all c.}$$

All firms must behave identically with respect to changes in transactions. To prove this condition, first notice that if $\beta_c = \overline{\beta}_c$ for all c, then $Cov(P_c \beta_c) = 0$. This occurs since,

$$Cov(P_c \beta_c) = \frac{1}{N} \sum_c (\beta_c - \overline{\beta}_c)(P_c - \frac{1}{N})$$

and implies that the macroparameter β will equal $\overline{\beta_c}$ since no aggregation bias exists. Formally,

$$\beta = \overline{\beta}_c + N\, Cov(P_c \beta_c) .$$

Since $\beta_c = \overline{\beta}_c = \beta$ and all three have constant values, $\Delta_1 M_i$ equals $\Delta_2 M_i$ as,

$$\Delta_1 M_i = \sum_c \beta_c \Delta t_{ci} = \overline{\beta}_c \sum_c \Delta t_{ci} = \beta \sum_c \Delta t_{ci} = \Delta_2 M_i .$$

No contradiction arises. Second, with no contradiction, distribution among firms of the aggregate change in transactions makes no difference. For example, if $\Delta t_{1i} = -\Delta t_{2i}$ and $\Delta t_{ci} = 0$ for all c greater than 2, then ΔT_i equals zero causing $\Delta_2 M_i$ to equal zero. Further, $\Delta_1 M_i = \sum_c \beta_c \Delta t_{ci} = (\beta_1 - \beta_2)\Delta t_{1i}$ which must equal zero since $\beta_1 = \beta_2$. In general, any change in transactions moves demand in the same proportion no matter who receives the additional business since all firms act homogeneously. Only in this case can aggregated data accurately represent a microrelation. When firms act differently, estimation must occur for each individual firm since the expectation of macroparameter does not equal its true value.

APPENDIX D

DATA

<u>Overview</u>. This Appendix gives detailed information on variables, sources, and uses of the samples employed in this dissertation.

<u>Variables</u>. The variables required for the models include: money balances (M), sales (S), accounts receivable (AR), accounts payable (AP), inventories (I), depreciation (D), cost of goods sold (COGS), gross property plant and equipment (GPPE), accumulated depreciation (AD), current assets (CA), gross investment (GI), treasury bill rate (TBR), federal funds rate (FFR), ten year corporate bond rate (CBR), equity returns (EQR), wholesale price index (WPI), product price index for particular industries (PPI), product price index for specific commodities (PPC), wage rates for specific commodities (WR), income tax (T), net income (NI), and repurchase agreement volume (RPV). The raw data provide for construction of additional relevant variables formed as follows.

Cash receipts (CR) appear as CR = S - ΔAT. Cash flows (CF) are constructed as CF = S - ΔAR - ΔI - COGS - ΔGPPE + ΔAP + ΔAP. Computing CF for each firm over the sample period allows forming of the variance of cash flow (V) for each firm. The amount of capital stock (K) appears as $K = GI + GPPE_{-1} - D$, where $GPPE_{-1}$ equals GPPE in the previous period. The price of capital becomes,

$$RR = \frac{(1 - uz)}{(1 - u)} \left(q(r + \delta) - \dot{q}\right)$$

where r equals the federal funds rate. q is the price of capital goods as measured by the PPC. δ stands for the rate of depreciation computed as D divided by K. u equals the tax rate as T divided by NI. Finally, z is the present value of the deduction for depreciation per dollar invested. Computation of z requires the following steps. First, the deduction is expressed as a percent of the value of capital for each time period. Second, those figures are discounted to obtain present value of the deductionper dollar invested in each period. Physical output (Q) is constructed as $Q = (S + \Delta I) \div PPC$. All real values equal nominal ones deflated by the WPI.

Sources. The bulk of data were obtained directly from the COMPUSTAT tapes which use the following primary sources.

Individual firms must file 10-K (annual) and 10-Q (quarterly) financial reports with the Securities and Exchange Commission. Also, firms make available both annual and quarterly reports to shareholders. Direct company contacts provide other specific information not generally available. Interactive Data Services Incorporated make stock data available. Additional variables are provided by various publications such as: *Dow Jones News Service*, Standard and Poor's *Stock Guides*, *Dividends Records*, and *Corporation Record's*. These sources give ovservations on M, S, AR, AP, I, GI, D, COGS, GPPE, AD, CA, T, NI, AND EQR, which are put onto the COMPUSTAT tapes.

The *Economic Report of the President Transmitted to Congress*, published by the U. S. Government Printing Office, contains observations on the following variables. The CBR, FFR, and TBR appear in Table B-67 while RPV is listed in Table B-62.

Wage rates for individual firms (WR) are contained ni the U. S. Department of Labor, Bureau of Labor Statistics, *Employment and Earnings*, Table C-2. This data exists for individual commodities by census code. By identifying its major product, a wage series can be constructed for each firm. For example, since General Motors would match with the motor vehicle entry, its wage rate would be found under that entry in Table C-2. When a firm has a diverse product line, the wage rate was formed as a weighted average of the wages associated with the different products. The weights come from analysis in The Value Line Investment Survey, published by Arnold Bernard and Company. For example, Graniteville Company receives about 70 percent of revenue from textiles and 30 percent from clothing retail trade. Its wage would equal 70 percent times the PPC for textiles plus 30 percent time the wage for clothing retail. This process is inferior to obtaining the information directly from each firm. Unfortunately, that data was not available.

The U. S. Department of Labor, Bureau of Labor Statistics, Producer Prices and Price Indexes, Tables 4 and 6 contain data on WPI, PPI, and PPC. The same matching between each firm and its major products allows the data on PPC to represent the price of a firm's output. In a similar manner, matching occurred between each firm and the major capital goods they purchase. The corresponding PPC or PPI then measures the price of capital goods (q) used to construct RR.

Uses. To allow for both estimation and prediction, each sample was split into two components. The quarterly one contains 30 observations per firm suggesting a trade off when selecting the size of each subsample. To maintain enough degrees of freedom for efficient estimation, the first 26 quarters go into the estimation subsample. The remaining four (about 16 percent) make up the prediction subsample allowing 380 forecasts. The annual data base, with 18 observations per firm, renders only two predictions per firm, leaving 16 for estimation. This split gives 384 predictions.

APPENDIX E

TRANSLOG COST MODEL

Overview. As Finnery (1980) showed, the transactions cost and neoclassical variants result in nearly identical money demand specifications. A general form of both the transactions cost approach and the neoclassical model comes from Saving (1972) and Nadiri (1969), respectively. In contrast, Dennis and Smith (1978) use a translog cost function to describe money's productive role. As Fisher (1974) argues, any derived money demand equation depends on the underlying functions used in the process. Thus, each procedure listed above represents a plausible embodiment of the production model. This appendix describes the translog cost process which renders the third specification.

Translog Cost Variant. The production process can be modeled by either the production function or its dual cost function. Using a cost function to estimate production relationships has several advantages to using the production function itself. First, a cost function is homogeneous of degree one in prices of inputs even if the production function does not exhibit the same property. This result stems from the fact that doubling all input prices must double a firm's costs. Yet, optimal factor combinations remain the same since relative input prices have not changed. Because of this property, generating estimates does not require imposing homogeneity conditions on the production process. Second, the estimating equations contain input prices as independent variables not factor quantities as with a production function. This result is desirable for two reasons. Theoretically, individual firms consider input prices as exogenous in their decision process. On the other hand, factor quantities are choice variables of profit maximization and, therefore, have stochastic elements. Empirically, input quantities tend to exhibit significant amounts of multicollinearity while factor prices do not. Third, using the production function method to obtain elasticity of substitution estimates requires inverting the matrix containing production function coefficient estimates--the bordered hessian. Such a procedure tends to exaggerate estimation error due to roundoff. Elasticity of substitution estimates do not require this inversion when using the cost function method. Finally, both procedures give equations which have linear logarithmic forms.

In general, the cost function approach describes the production relationships in terms of Allen partial elasticities of substitution. Corresponding to the cost minimization problem,

$$\text{MIN } C = \sum_{h=1}^{n} X_h P_h$$

$$\text{Subject to} \quad Q = F(X_1 \ . \ . \ . \ X_n)$$

where X_h represents input amounts, P_h stands for input prices, and Q is output. There exists a dual minimum cost function as,

$$C^* = g(Q, P_1 \ . \ . \ . \ P_n).$$

This function relates, to every combination of the P_h , the minimum cost (C^*) corresponding to the optimum--profit maximizing--levels of $X_h(\overset{*}{X_h})$. In other words, C refers to the production cost of any feasible factor combination while C* gives the expense of cost minimizing input combinations. Thus, C* corresponds to the firm's expansion path and is a function of factor prices since they determine optimum input combinations. Finally, the minimum cost function is homogeneous of degree one in factor prices regardless of the homogeneity properties of the production function. The importance of C* lies in the basis it provides for deriving both Allen partial and factor demand elasticies.

Specifically, Shepard's (1953) lemma shows that the derived factor demand equals the partial derivative of C* with respect to factor prices as,

$$\frac{2C^*}{2P_i} = X_i \; .$$

Furthermore, if f symbolizes the bordered hessian of the production function and f_i, f_{ij} are the partial and cross partial derivatives, respectively, then the Allen partial elasticity of substitution for two inputs i and j equal,

$$\sigma_{ij} = \frac{\sum_h^n X_h f_h}{X_i X_j} (f^{-1})_{ij}$$

where $(f^{-1})_{ij}$ stands for the ij^{th} element of f^{-1}. Because of the symmetry of f^{-1}, σ_{ij} equals σ_{ji}. Furthermore, cost functions can compute estimates of σ_{ij}, without a matrix inversion, directly from the parameters of the cost function as,

$$\sigma_{ij} = \frac{\sum_h^n P_h X_h}{X_i X_j} \frac{\partial^{\partial} C^*}{\partial P_i \partial P_j}$$

which can be transformed, by dividing and multiplying the right side by $\frac{P_j}{X_i}$, into,

$$\sigma_{ji} = \sigma_{ij} = \frac{\eta_{ij}}{C_j}$$

where

$$\eta_{ij} = \frac{2X_i}{2P_j}\frac{P_j}{X_i}$$

$$C_j = \frac{X_j P_j}{\sum_h^n X_h P_h}.$$

Here η_{ij} represents the cross elasticity of demand for X_i with respect to P_j and C_j equals the relative cost share of X_j.

The proof of these results comes from the Lagrangian form for the cost minimization problem as,

$$L = \sum_h^n X_h P_h + \lambda(F(X_1 X_2 \ldots Xn) - Q)$$

which has as first order conditions,

$$P_h - \lambda f_h = 0$$

$$f(X_1 \ldots X_n) - Q = 0.$$

Expanding to obtain the total differential of these conditions renders the following system.

$$\lambda \begin{bmatrix} 0 & f_1 & \cdot & \cdot & \cdot & f_n \\ f_1 & f_{11} & \cdot & \cdot & \cdot & f_{1n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ f_n & f_{n1} & \cdot & \cdot & \cdot & f_{nn} \end{bmatrix} \begin{bmatrix} d\lambda/\lambda \\ dX_1 \\ \cdot \\ \cdot \\ dX_n \end{bmatrix} = \begin{bmatrix} \lambda dQ \\ dP_1 \\ \cdot \\ \cdot \\ dP_n \end{bmatrix}$$

Solving the system for the endogenous variables gives,

$$\begin{bmatrix} dLNX \\ dX_1 \\ \cdot \\ \cdot \\ dX_n \end{bmatrix} \quad \begin{bmatrix} dLN\lambda \\ dX_1 \\ \cdot \\ \cdot \\ dX_n \end{bmatrix} = \frac{1}{\lambda} f^{-1} \begin{bmatrix} dQ \\ dP_1 \\ \cdot \\ \cdot \\ dP_n \end{bmatrix}$$

where f^{-1} stands for the inverse of the matrix of partial derivatives. This result implies that,

$$\frac{\partial X_i}{\partial P_j} = \frac{1}{\lambda} (f^{-1})_{ij}$$

where $(f^{-1})_{ij}$ is the ij^{th} element of f^{-1}. Substitution of this condition into the definition for σ_{ij} gives,

$$\sigma_{ij} = \frac{\sum_h^n X_h f_h}{X_i X_j} \lambda \frac{\partial X_i}{\partial P_j} .$$

Also solving the first order conditions for f_i gives,

$$f_i = \frac{P_i}{\lambda}$$

which, by substitution, makes the elasticity

$$\sigma_{ij} = \frac{\frac{1}{\lambda} \sum_h X_h P_h}{X_i X_j} \lambda \frac{\partial X_i}{\partial P_j} .$$

Finally, since $\frac{\partial C^*}{\partial P_i} = X_i$,

$$\frac{\partial X_i}{\partial P_j} = \frac{\partial^2 C^*}{\partial P_i \partial P_j}$$

which means the elasticity of substitution equals,

$$\sigma_{ij} = \frac{\Sigma\ X_h\ P_h}{X_i\ X_j}\frac{\partial^2 C^*}{\partial P_i \partial P_j}$$

multiplying and dividing the right side by $\frac{P_j}{X_i}$,

$$\sigma_{ij} = \frac{X_i \Sigma\ X_h\ P_h}{P_j\ X_i\ X_j}\ \frac{\partial X_i}{\partial P_j}\frac{P_j}{X_i}$$

$$\sigma_{ij} = \frac{\Sigma\ X_h\ P_h}{X_j\ X_j} \cdot \frac{\partial X_i}{\partial P_j}\frac{P_j}{X_i}$$

Since $\eta_{ij} = \frac{\partial X_i}{\partial P_j}\frac{P_j}{X_i}$ and $C_j = \frac{X_j P_j}{\Sigma\ X_h P_h}$,

$$\sigma_{ij} = \frac{\eta_{ij}}{C_j}$$

After estimating parameters of the specific form of C*, the last equation gives a method to compute elasticities of substitution for given amounts of both factors and total cost. This process works well, specifically, with a translog cost function.

The translog cost function is a logarithmic Taylor series expansion, to the second degree, of a twice differentiable analytic cost function around values of zero for both LN(Q) and LN(P_i). C* in logarithmic form appears as,

$$LN\ C^* = f\big(LN(Q)\ LN(P_1)\ .\ .\ .\ LN(P_n)\big)$$

and has derivates at zero equal to,

$$LN(C^*) = \alpha_0 ; \frac{\partial\ LN(C^*)}{\partial\ LN(Q)} = \alpha_Q ;$$

$$\frac{\partial\ LN(C^*)}{\partial\ LN(P_i)} = \beta_i ; \frac{\partial^2\ LN(C^*)}{\partial\ LN(P_i)\ \partial\ LN(P_j)} = \gamma_{ij} ;$$

$$\frac{\partial^2\ LN(C^*)}{\partial\ LN(P_i)\ \partial\ LN(Q)} = \delta_i .$$

Since cross derivatives are equal, the constraint $\gamma_{ij} = \gamma_{ji}$ must hold true. The resulting Taylor series expansion appears as,

$$LN(C) = \alpha_0 + \alpha_Q\ LN(Q) + \frac{1}{2}\,\alpha_{QQ}\ \big(LN(Q)\big)^2 + \sum_{i=1} \beta_i\ LN(P_i) +$$

$$\frac{1}{2} \sum_i \sum_j \gamma_{ij}\ LN(P_i)\ LN(P_j) + \sum_{i=1} \delta_i\ LN(P_i)\ LN(Q)$$

which represents a functional form, if the remainder is ignored and all derivates remain constant. The latter condition always holds for parameters of a regression equation. Also, the property of homogeneity in factor prices implies;

$$\sum_i \beta_i = 1$$

$$\sum_i \gamma_{ij} = \sum_j \gamma_{ji} = \sum_i \sum_j \gamma_{ij} = 0$$

$$\sum_i \delta_i = 0.$$

Yet, it does not impose any homogeneity constraints on the production function. In fact, almost no restrictions are put on elasticities of substitution or factor demands making the translog model more general than commonly used alternatives.

Estimation of the function can occur directly or in its first derivatives which constitute factor shares of total cost (C_i).

$$\frac{\partial\ LN(C^*)}{\partial\ LN(P_i)} = \frac{\partial C^*}{\partial P_i}\ \frac{P_i}{C^*} = X_i^*\ \frac{P_i}{C^*} = C_i.$$

For the translog function C_i equals,

$$\frac{\partial\ LN(C^*)}{\partial\ LN(P_i)} = \beta_i + \sum_j \gamma_{ij}\ LN(P_j) + \delta_i\ LN(Q)$$

The γ_{ij} have little importance themselves. However, they can construct elasticities of both substitution and factor demand as follows,

$$\sigma_{ij} = \frac{\gamma_{ij}}{C_i C_j} + 1$$

$$\sigma_{ii} = \frac{\gamma_{ij} + C_i^2 - C_i}{C_i^2}$$

$$\eta_{ij} = \frac{\gamma_{ij}}{C_i} + C_i$$

$$\eta_{ij} = \frac{\gamma_{ij}}{C_i} + C_i - 1.$$

Proving these results starts from the elasticity of substitution as,

$$\sigma_{ij} = \frac{\sum_h X_h P_h}{X_i X_j} \frac{\partial^2 C^*}{\partial P_i \partial P_j}$$

as derived earlier. Furthermore, the γ_{ij} coefficient of the translog function is,

$$\gamma_{ij} = \frac{\partial^2 \; LN(C^*)}{\partial \; LN(P_i) \; \partial \; LN(P_j)} .$$

or in long form

$$\gamma_{ij} = \frac{\partial\left(\frac{\partial \; LN(C^*)}{\partial \; LN(P_i)}\right)}{\partial \; LN(P_j)}$$

which expands to,

$$\gamma_{ij} = \frac{\partial\left(\frac{\partial C^*}{\partial P_i} \frac{P_i}{C^*}\right)}{\partial \; LN(P_j)}$$

Since $\partial \; LN(P_j) = \frac{\partial P_j}{P_j}$,

$$\gamma_{ij} = \frac{P_j \partial\left(\frac{\partial C^*}{\partial P_i} \frac{P_i}{C^*}\right)}{\partial P_j}$$

Carrying out the differentiation by parts gives,

$$\gamma_{ij} = P_j\left(\frac{\partial^2 C^*}{\partial P_i \partial P_j} \cdot \frac{P_i}{C^*} - \frac{P_i}{(C^*)^2} \frac{\partial C^*}{\partial P_j} \cdot \frac{\partial C^*}{\partial P_i}\right).$$

Because $\frac{\partial C^*}{\partial P_h} = X_h^*$,

$$\gamma_{ij} = \frac{P_i P_j}{C^*} \frac{\partial^2 C^*}{\partial P_i \partial P_j} - \frac{P_i P_j}{(C^*)^2} X_i X_j$$

Yet, $\frac{P_i X_i}{C^*} = C_i$ and $\frac{P_j X_j}{C^*} = C_j$.

Thus,

$$\gamma_{ij} = \frac{P_i P_j}{C^*} \frac{\partial^2 C^*}{\partial P_i P_j} - C_i C_j \ .$$

Solving this equation for the second partial gives,

$$\frac{\partial^2 C^*}{\partial P_i \partial P_j} = \frac{C^*}{P_i P_j} (\gamma_{ij} + C_i C_j)$$

Substituting this expression into the formula for σ_{ij} renders,

$$\sigma_{ij} = \frac{\sum_h^n P_h X_h}{X_i X_j} [\frac{C^*}{P_i P_j} (\gamma_{ij} + C_i C_j)]$$

$$\sigma_{ij} = \frac{(\Sigma \ P_h X_h) C^*}{P_i X_i \ P_j X_j} (\gamma_{ij} + C_i C_j).$$

Since $C_i = \frac{P_i X_i}{C^*}$ and $C_j = \frac{P_j X_j}{\Sigma P_i X_i}$,

$$\sigma_{ij} = \frac{1}{C_i C_j} (\gamma_{ij} + C_i C_j)$$

$$\sigma_{ij} = \frac{\gamma_{ij}}{C_i C_j} + 1$$

The proof for σ_{ii} is similar except $\frac{\partial P_i}{\partial P_i} = 1$ which gives $-C_i$, not one.

The proof of η_{ij} and η_{ii} follow from the relationship derived as,

$$\sigma_{ij} = \sigma_{ji} = \frac{\eta_{ij}}{C_j}.$$

Thus,

$$\eta_{ij} = \sigma_{ij} C_j.$$

By substitution,

$$\eta_{ij} = (\frac{\gamma_{ij}}{C_i C_j} + 1)C_j$$

$$\eta_{ij} = \frac{\gamma_{ij}}{C_i} + C_j$$

and

$$\eta_{ii} = \sigma_{ii} C_i.$$

By substitution,

$$\eta_{ii} = \frac{1}{C_i^2} (\gamma_{ii} + C_i^2 - C_i)C_i$$

$$\eta_{ii} = \frac{\gamma_{ii}}{C_i} + C_i - 1.$$

Once the γ_{ij} are estimated, they combine with known factor cost shares allowing the formulas to compute all relevant elasticities. Furthermore, since such estimated elasticities are linear transformations of the γ_{ij}, which have known econometric properties, the corresponding properties for the elasticities follow directly. This conclusion provides a basis for statistical testing of the function's underlying properties. One such characteristic of special importance is separability.

The translog cost function does not restrict the production process to a homothetic one where inputs are separable. This generality, allows tests to validate or refute this condition. If the restrictions prove true, the translog model reduces to a simpler form. When the constraints are not upheld, estimates should use the translog form. The test for input separability relies on the following analysis.

Consider a twice differentiable production function as,

$$Q = F(X_1 \; . \; . \; . \; X_n)$$

where the set of n inputs can partition into r mutually exclusive subsets $N_1 \; . \; . \; . \; N_r$. By definition, the production function exhibits weak separability to a partition if the marginal rate of substitution between two inputs, i and j, within the subset is independent of input quantities outside the subset (say k). Since,

$$MRS_{ij} = \frac{F_i}{F_j}$$

Q is weakly separable when,

$$\frac{\partial \left(\frac{F_i}{F_j}\right)}{\partial X_k} = 0.$$

Expanding gives,

$$F_j F_{ik} - F_i F_{jk} = 0$$

which makes $Q = F(X_1 \ldots X_k, (X_i X_j))$. This result has the following intuitive explanation.

Holding the use of i and j constant, if an increase in k causes equal increases in the efficiency of i and j, then the latter two are weakly separable from k. In other words, the additional k shifts the marginal product schedules of i and j proportionally. k holds either an identical substitute or complimentary relation to both i and j. Thus, the elasticity of substitution between i and k will equal that of j and k. This property allows the firm to treat k separately from i and j in the profit maximization process. Optimal factor ratios can be set first within each subset, independently of other inputs. Then optimal combinations are fixed between subsets holding within subset ratios constant. These properties extend to the translog cost function as follows.

Dual to the production function exists a cost funcion as,

$$C^* = c(Q, P_1 \ldots P_n)$$

which is homothetic when C* can be re-written as,

$$C^* = H(Q) \cdot G(P_1 \ldots P_n)$$

where H and G represent functions of only output and only prices, respectively. When G has nonzero first and second derivatives, weak separability of the i and j input prices from that of k requires,

$$G_j G_{ik} - G_i G_{jk} = 0$$

Furthermore, if G exhibits such separability, the production function must have the same characteristic. This duality implies that $\sigma_{ik} = \sigma_{jk}$, which gives an integral part to the proof of the above equation.

The Allen partial elasticities appear as,

$$\sigma_{ik} = \frac{\sum_{h=1}^{n} P_h X_h}{X_i X_k} \frac{\partial X_i}{\partial P_k}$$

$$\sigma_{jk} = \frac{\sum_{h=1}^{n} P_h X_h}{X_j X_k} \frac{\partial X_j}{\partial P_k}$$

Also,

$$X_i = \frac{\partial C}{\partial P_i} = H(Q)G_i$$

$$X_j = \frac{\partial C}{\partial P_j} = H(Q)G_j$$

$$X_k = \frac{\partial C}{\partial P_k} = H(Q)G_k.$$

Therefore,

$$\frac{\partial X_i}{\partial P_k} = H(Q)G_{ik}$$

$$\frac{\partial X_j}{\partial P_k} = H(Q)G_{jk}.$$

Remembering that $C^* = H(Q) \cdot G(P)$, substitution gives,

$$\sigma_{ik} = \frac{H(Q) \cdot G(P)}{H(Q)G_i \cdot H(Q)G_k} H(Q)G_{ik}$$

$$\sigma_{ik} = \frac{G\ G_{ik}}{G_i G_k}$$

and similarly,

$$\sigma_{jk} = \frac{G\ G_{jk}}{G_j G_k} .$$

Separability means $\sigma_{ik} = \sigma_{jk}$ which occurs only if $G_j G_{ik} - G_i G_{jk} = 0$. For the translog model, this condition can be written in terms of cost shares and parameters.

Since,

$$G_i = \frac{\partial\ LN(C)}{\partial\ LN(P_i)} = C_i$$

$$G_{jk} = \frac{\partial^2\ LN(C)}{\partial\ LN(P_j)2\ LN(P_k)} = \gamma_{ij}$$

$$G_j = \frac{\partial\ LN(C)}{\partial\ LN(P_j)} = C_j$$

$$G_{ik} = \frac{\partial^2\ LN(C)}{\partial\ LN(P_i)2\ LN(P_k)} = \gamma_{jk}.$$

Separability means,

$$C_j \gamma_{ik} - C_i \gamma_{jk} = 0.$$

Since C_j and C_i exceed zero this condition holds linearly if $\gamma_{ik} = \gamma_{jk} = 0$. It holds for a non-linear form as,

$$\gamma_{kk} = \frac{\gamma_{kk}^2}{\gamma_{jj}} \qquad \alpha_k = 1 + \frac{\beta_j \gamma_{jk}}{\gamma_{jj}} .$$

The formation of these underlying properties and the elasticities require estimates of the cost function parameters.

Estimation of the translog coefficients can occur indirectly by taking first derivatives of LN(C) as,

$$\frac{\partial \, LN(C)}{\partial \, LN(P_i)} = \frac{\partial C}{\partial P_i} \frac{P_i}{C} = \frac{P_i X_i^*}{C} = \beta_i + \sum_j^3 \gamma_{ij} \, LN(P_i) + \delta_i \, LN(Q)$$

where X_i^* stands for the cost minimum amount of input i. The first expression equals the partial derivative of the cost function with respect to the logarithm of the price for the i^{th} input. Rewriting renders the second term which represents the elasticity of costs with respect to the price of the i^{th} input. Due to profit maximum behavior, the expression equals the relative cost share of the i^{th} input (C_i). For the three-input case under study, the costs shares of labor, money, and capital appear as,

$$C_l = \beta_1 + \gamma_{11} \, LN(P_l) + \gamma_{12} \, LN(P_m) + \gamma_{13} \, LN(P_k) + \delta_1 \, LN(Q) + \varepsilon_1$$

$$C_m = \beta_2 + \gamma_{21} \, LN(P_l) + \gamma_{22} \, LN(P_m) + \gamma_{23} \, LN(P_k) + \delta_2 \, LN(Q) + \varepsilon_2$$

$$C_k = \beta_3 + \gamma_{31} \, LN(P_l) + \gamma_{32} \, LN(P_m) + \gamma_{33} \, LN(P_k) + \delta_3 \, LN(Q) + \varepsilon_3$$

The linear homogeneity property also implies that parameter values of any two equations exactly identify all parameters of the system. Furthermore, unrestricted estimation of each remaining equation does not ensure that $\gamma_{ij} = \gamma_{ji}$. Thus, after omitting C_k, estimation requires that the following restrictions be substituted into the first two equations.

$$\gamma_{13} = \gamma_{31} = (-\gamma_{11} - \gamma_{21})$$

$$\gamma_{23} = \gamma_{32} = (-\gamma_{22} - \gamma_{21})$$

Solving gives the estimating equation as,

$$\begin{bmatrix} C_1 \\ \\ C_m \end{bmatrix} = \begin{bmatrix} 1 & 0 & LN\frac{P_1}{P_k} & LN\frac{P_m}{P_k} & 0 & LN(Q) & 0 \\ \\ 0 & 1 & 0 & LN\frac{P_1}{P_k} & LN\frac{P_m}{P_k} & 0 & LN(Q) \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \gamma_{11} \\ \gamma_{21} \\ \gamma_{22} \\ \delta_1 \\ \delta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}$$

which represents a testable form. Thus, three equations representing the production approach exist. One reflects the transactions cost model, another embodies the neoclassical theory, and the third employs a translog cost specification.

APPENDIX F

ESTIMATION OF THE TRANSLOG COST MODEL

Overview. This appendix presents the process required to estimate and test the translog cost model. For details of this function, see Appendix E.

Estimation. The translog cost model uses three cost share equations as,

$$C_m = \alpha_m + \gamma_{mm}LN(FFR) + \gamma_{ml}LN(WR) + \gamma_{mk}LN(RR) + \delta_m LN(Q) + \varepsilon_m \qquad (4\text{-}32)$$

$$C_l = \alpha_l + \gamma_{lm}LN(FFR) + \gamma_{ll}LN(WR) + \gamma_{lk}LN(RR) + \delta_l LN(Q) + \varepsilon_l \qquad (4\text{-}33)$$

$$C_k = \alpha_k + \gamma_{km}LN(FFR) + \gamma_{kl}LN(WR) + \gamma_{kk}LN(RR) + \delta_k LN(Q) + \varepsilon_k \qquad (4\text{-}34)$$

where the three disturbance terms arise due to errors in carrying out profit maximizing outcomes. This property introduces one of two complexities outlined below.

Because these errors occur in all input markets, the three disturbance terms will be correlated. Thus, the set of cost shares poses the problem of seemingly unrelated regressions. Efficient estimation necessitates the use of two-stage Zellner (2SZ) method.

In contrast, the errors may, in part, determine the values of output as well as the disturbances. This result stems from the fact that firms treat output amounts as a choice variable in the profit maximization process. In this case, the cost share model represents a truly simultaneous set since an explanatory variable of one equation has correlation with the disturbances of the remaining equations. Efficient estimation requires a systems method, such as three stage least squares (3SLS), to account for the cross equation correlation. Both 2SZ and 3SLS will generate estimates for the cost share set.

Such estimation must account for the following properties of the system. First, at every observation the C_i must sum to unity implying that the corresponding ε_i sum to zero. Second, the translog cost function has first degree homogeneity in all input prices. Doubling factor prices causes costs to double since relative prices and, thus, factor usage stays the same. These characteristics imply the following restrictions.

$$\sum_i \alpha_i = 1$$

$$\sum_i \gamma_{ij} = \sum_j \gamma_{ji} = \sum_i \sum_j \gamma_{ij} = 0$$

$$\sum_i \delta_i = 0$$

where $i \neq j$ and both sum over m, l, and k. In addition, the parameters of any two cost share equations just identify the coefficients of the whole set implying that only two equations can undergo simultaneous estimation. Unfortunately, both 2SZ and 3SLS generate estimates sensitive to which equation is omitted. As shown by Dhrymes (1973), iteration of the 2SZ or 3SLS methods results in estimates lacking such sensitivity. The iterative two stage Zellner (I2SZ) estimators will give consistent and asymptotically efficient results, if cross equation correlation exists only between disturbances. When explanatory variables have correlation with disturbances of other equations, iterative three stage least squares (I3SLS) provides consistent estimators with improved asymptotic efficiency compared to those of I2SZ. These two methods generate estimates for the cost shares of money balances and labor which, in turn, construct estimates for the omitted equation (capital). Before applying I2SZ and I3SLS, the exact form of the system must be discovered.

In particular, symmetry constraints, implied by the Taylor series expansion which creates the translog function, take the form,

$$\gamma_{ij} = \gamma_{ji} \qquad \text{for all } i \neq j$$

since only C_m and C_l remain in the cost share set these restrictions reduce to,

$$\gamma_{km} = \gamma_{mk} = -(\gamma_{mm} + \gamma_{lm})$$

$$\gamma_{kl} = \gamma_{lk} = -(\gamma_{ll} + \gamma_{lm}) .$$

Testing their validity requires estimation of both an unconstrained system as Equations (4-32) and (4-33) and a constrained version. The latter imposes the symmetry restrictions on the unconstrained set to form,

$$\begin{bmatrix} C_m \\ C_l \end{bmatrix} = \begin{bmatrix} 1 & 0 & LN(\frac{FFR}{RR}) & LN(\frac{WR}{RR}) & 0 & LN(Q) & 0 \\ 0 & 1 & 0 & LN(\frac{FFR}{RR}) & LN(\frac{WR}{RR}) & 0 & LN(Q) \end{bmatrix} \begin{bmatrix} \alpha_m \\ \alpha_l \\ \gamma_{mm} \\ \gamma_{ml} \\ \gamma_{ll} \\ \delta m \\ \delta_l \end{bmatrix} + \begin{bmatrix} \varepsilon_m \\ \varepsilon_l \end{bmatrix} \qquad (4\text{-}35)$$

Given this equation, estimates of the parameters construct the underlying production relationships.

The Allen partial elasticity for inputs i and j(σ_{ij}) measures the impact on the usage of i of a change in the price of j holding all other things constant. Construction from the coefficients of Equation (4-35) occurs as,

$$\sigma_{ij} = \frac{\gamma_{ij}}{\overline{C}_i \overline{C}_j} + 1$$

$$\sigma_{ii} = \frac{1}{\overline{C}_i^{\,2}} (\gamma_{ii} + \overline{C}_i^2 - \overline{C}_i)$$

where $\overline{C}_i$ stands for the mean value of the i^{th} input's cost share and measures the cost share's central tendency. Assuming $\overline{C}_i$ nonstochastic allows a simple calculation of the standard error of σ_{ij} as,

$$SE(\sigma_{ij}) = \frac{SE(\gamma_{ij})}{\overline{C}_i \overline{C}_j} .$$

Since a random process actually generates C_i and C_j only the asymptotic properties of these estimates are known.

In a similar manner, the derived factor demand elasticity of i with respect to j(η_{ij}) measures the change in usage of input i due to a change in the price of input j, other things equal. Construction from Equation (4-35) coefficients occurs as,

$$\eta_{ij} = \frac{\gamma_{ij}}{\overline{C}_i} + \overline{C}_j$$

$$\eta_{ii} = \frac{\gamma_{ii}}{\overline{C}_i} + \overline{C}_i - 1$$

with standard error of,

$$SE(\eta_{ij}) = \frac{SE(\gamma_{ij})}{\overline{C}_i} .$$

Tests. In addition to providing these elasticities, the translog cost model also renders information about underlying production structures. In particular, the production function is well behaved if it increases monotonically with an increase in all inputs and has convex isoquants. In addition to these two properties, input separability can be tested. First, Berndt and Christensen (1973a) have shown that monotonicity requires that all cost share values fitted by estimation exceed zero at every observation. Inspection showed that the I3SLS fitted values satisfy this criterion implying that output increases monotonically with an increase in all inputs. Second, strict convexity of isoquants necessitates negative definite bordered hessians at every observation as proved by Katzner (1970). Construction gives only negative definite hessians which means the production function has strictly convex isoquants.

Finally, separability represents a property of particular importance. It establishes the validity of describing the influence of the separable inputs independently of other inputs. For example, when capital (K) and labor (L) are separable from money (M), a production function omitting money balances can estimate the effect of K and L on output. In contrast, absence of separability requires the presence of all significant inputs in the production function to make unbiased estimates of any one's marginal product. In general, separability occurs in various degrees of severity.

The strongest degree, termed global separability, occurs when the cross partial derivatives of the translog cost function with respect to all pairs of input prices $(\frac{\partial^2 C^*}{\partial P_i \partial P_j})$ equal zero. This condition means that each input's effect on output is independent of the usage of all other inputs. For the translog cost model, the cross partial derivative above equals γ_{ij} and, thus, restrictions imposed by global separability appear as,

$$\gamma_{ml} = \gamma_{mk} = \gamma_{lk} = 0 .$$

Testing these restrictions requires estimating an unconstrained and constrained regression and computing the resulting f statistic. The calculations result in a statistic which exceeds its critical value and, thus, rejects global separability. The underlying production structure must allow for cross effects as embodied in the γ_{ij}. Yet, a less strigent form of separability may exist.

Weak separability can occur in three forms for the translog cost model under study; money and labor from capital (ML-K), money and capital from labor (MK-L), or labor and capital from money (LK-M). Weak separability implies that the pair of inputs influence output independently of the third input. Formally, the partial derivative of the marginal rate of substitution between the pair of inputs with respect to the third input $(\frac{\partial MRS_{ml}}{\partial K})$ equals zero.

As shown by Berndt and Christensen (1973a), the presence of each form of weak separability renders the following restrictions.

$$\text{ML-K} \quad C_m\gamma_{lk} - C_l\gamma_{mk} = 0$$

$$\text{MK-L} \quad C_m\gamma_{lk} - C_k\gamma_{ml} = 0$$

$$\text{LK-M} \quad C_l\gamma_{mk} - C_k\gamma_{ml} = 0$$

Since all C_i exceed zero, these conditions hold linearly as,

$$\text{ML-K} \quad \text{IF} \quad \gamma_{lk} = \gamma_{mk} = 0$$

$$\text{MK-L} \quad \text{IF} \quad \gamma_{lk} = \gamma_{ml} = 0$$

$$\text{LK-M} \quad \text{IF} \quad \gamma_{mk} = \gamma_{ml} = 0$$

or they exist in non-linear form as,

$$\text{ML-K} \quad \text{IF} \quad \gamma_{kk} = \frac{\gamma_{kk}^2}{\gamma_{ll}} \quad \text{AND} \quad \alpha_k = 1 + \frac{\alpha_l\gamma_{lk}}{\gamma_{ll}}$$

$$\text{MK-L} \quad \text{IF} \quad \gamma_{kk} = \frac{\gamma_{kk}^2}{\gamma_{ll}} \quad \text{AND} \quad \alpha_k = \frac{(\alpha_l - 1)\ \gamma_{lk}}{\gamma_{ll}}$$

$$\text{LK-M} \quad \text{IF} \quad \gamma_{kk} = \frac{\gamma_{lk}^2}{\gamma_{ll}} \quad \text{AND} \quad \alpha_k = \frac{\alpha_l\gamma_{lk}}{\gamma_{ll}}.$$

Testing for these six types of weak separability requires estimating both an unconstrained and constrained regression in each case. The regression results construct an f statistic which measures the additional explanatory power when lifting the restriction. Thus, if the calculated f exceeds its critical value, the corresponding type of weak separability is rejected. These tests, along with computations of elasticities, are carried out in Chapter IV.

APPENDIX G

TEST PROCEDURES

Overview. Three tests appear in Chapter IV which have special significance when using pooled data. They are outlined below.

Bartlett's Test. Bartlett's test has two properties which make it desirable when using pooled data. First, it does not involve a specific assumption about the nature of heteroscedasticity; thus, a general correction can occur, if necessary. Second, pooling creates groups which the test does require. Specifically, the procedure applies under the following conditions.

The observations available for estimation equal N in number and split into G cross-sectional groups. Thus, each group contains N_i unique observations. Finally, the mean value for the dependent variable in the g^{th} group equal $\overline{Y}_g$.

The test takes the form,

$$H_0: \ \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_g^2$$

$$H_a: \ \text{Not } H_0$$

and is conducted as follows. First, calcualte a consistent estimate of each σ_g^2 as,

$$S_g^2 = \frac{1}{Ng} \sum_i^{Ng} (Y_i - \overline{Y}_g)^2.$$

Second, use these estimates to create a test statistic equal to,

$$S = - \frac{N \ LN(\sum_g^G \frac{Ng}{N} S_g^2) - \sum_g^G N_g \ LN(S_g^2)}{1 + \frac{1}{3(G-1)} (\sum_g^G \frac{1}{Ng} - \frac{1}{N})}$$

Under the null hypothesis, the statistic follows a Chi-square distribution with G-1 degrees of freedom. The null hypothesis is rejected if S calculated exceeds the tabulated value at the chosen level of significance. If rejection occurs, the proper adjustment divides all observations by a consistent estimate of σ_i.

Adjusted Durbin-Watson Test. The adjusted Durbin-Watson test has the hypothesis structure

$$H_0: \rho_i = 0$$

$$H_1: \text{Not } H_0$$

where ρ_i symbolizes the coefficient of autoregression for each cross-section.

The test involves the calculation of a statistic using the residuals from OLS estimation. For each group, the statistic becomes

$$D_g = \frac{\sum_{t=2}^{Ng} (e_t - e_{t-1})^2}{\sum_{t=2}^{Ng} e_t^2}$$

where Ng represents the number of observations for the g^{th} group. The final statistic for all groups appears as,

$$D = \frac{\sum_{g=1}^{G} D_j}{G}$$

when successive values of e_t for every group lie close together, D will be small indicating the presence of positive serial correlation. In general, D lies between zero and four with a value of two indicating absence of serial correlation.

In a sample of data, positive serial correlation is rejected if D lies below a lower bound (d_L). If D exceeds an upper bound (d_U), the null hypothesis cannot be rejected. When D falls between d_L and d_U, prudence demands rejecting the null hypothesis. Since the distribution of D, in this case, is unknown this test merely suggests the likelihood of the presence or absence of autocorrelation. It does not give a precise measure.

Adjusted Von-Neuman Ratio Test. The Von-Neuman ratio test has the hypothesis structure,

$$H_0: \ \rho_i = 0$$

$$H_1: \ \text{Not } H_0$$

where ρ_i stands for the coefficient of autocorrelation for the i^{th} firm.

The test employs the ordinary least squares residuals (E) to form a statistic for each firm as,

$$q_i = \frac{(n-1)^{-1} \sum_{t=2}^{n} (e_t - e_{t-1})^2}{n^{-1} \sum_{t=1}^{n} (e_t - \bar{e})^2}$$

where n represents the number of predictions for the i^{th} firm and $\bar{e}$ stands for the mean value of the prediction error series. The test statistic then averages these ratios over all firms to give,

$$Q = \frac{\sum_{i}^{G} q_i}{G} .$$

This procedure parallels the one used in the adjusted Durbin-Watson test. If the calculated value of Q lies sufficiently below two (as given by a critical value), first order autocorrelation exists.

Since the distribution of Q is not known, this test does not provide precise results. It only suggests the existance or absence of autocorrelation.

Zero Correlation Coefficient Test. Consider the least square residuals from Equations (4-1) and (4-3) as $e_i^{(1)}$ and $e_i^{(3)}$ for the i^{th} observation. Suppose that $(e_i^{(1)}, e_i^{(3)})$, $i = 1 \ldots N$, makes a random sample from a bivariate normal distribution with zero means, variances of σ_1^2 and σ_3^2, and correlation coefficient ρ_{13}. The pair of random variables $e^{(1)} + e^{(3)}$ and $e^{(1)} - e^{(3)}$ will have a covariance as,

$$COV = E[(e^{(1)} + e^{(3)})\ (e^{(1)} - e^{(3)})] + E(e^{(1)} + e^{(3)})\ E(e^{(1)} - e^{(3)}).$$

Because both residual series are unbiased,

$$COV = E[(e^{(1)} + e^{(3)})\ (e^{(1)} - e^{(3)})]$$

$$COV = E[(e^{(1)}e^{(1)} - e^{(3)}e^{(3)})]$$

$$COV = \sigma_1^2 - \sigma_3^2$$

which implies that the two variances are equal (thus, $S_1^2 = S_3^2$) if and only if the two random variables have zero correlation. The sample estimate of the correlation coefficient employs the least square residuals as,

$$r_{13} = \frac{\sum_i (e_i^{(1)} + e_i^{(3)})\,(e_i^{(1)} - e_i^{(3)})}{[\sum_i (e_i^{(1)} + e_i^{(3)})^2 \sum_i (e_i^{(1)} - e_i^{(3)})^2]^{1/2}} \cdot$$

The test has the null hypothesis that the two variances are equal and employs a t statistic, formed as r divided by its standard error, under the assumption that r has a normal distribution. Fisher (1958) has shown this assumption to hold in sample sizes exceeding 100 observations. Rejection of the null hypothesis occurs when the magnitude of the calculated statistic exceeds the critical value.

VITA

Jeffrey Merrill Herbener

Candidate for the Degree of

Doctor of Philosophy

Thesis: THE DEMAND FOR MONEY BY FIRMS

Major Field: Economics

Biographical:

Personal Data: Born in Canton, Illinois, October 23, 1955, the son of Lloyd and Ann Herbener.

Education: Graduated from Wahoo High School, Wahoo, Nebraska, in May 1973; earned a Bachelor of Science degree in Economics from Nebraska Wesleyan University in May 1977; received a Master of Science degree in Economics from Oklahoma State University in December 1981; completed the requirements for the Doctor of Philosophy degree at Oklahoma State University in December 1982.

Professional Experience: Graduate teaching associate, Oklahoma State University, 1979-1982; Instructor, Metropolitan Technical Community College, Summer 1980; Assistant Professor, Pittsburg State University, 1982.