Title
Debt financing and the dynamics of agency costs

Permalink
https://escholarship.org/uc/item/0m68d61d

Author
Cao, Bolong

Publication Date
2006

Peer reviewed|Thesis/dissertation
Debt Financing and the Dynamics of Agency Costs

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Economics

by

Balong Cao

Committee in charge:

Professor Garey Ramey, Chair
Professor Takeo Hoshi
Professor Bruce Lehmann
Professor Yixiao Sun
Professor Michelle J. White

2006
The dissertation of Bolong Cao is approved, and it is acceptable in quality and form for publication on microfilm:

Chair

University of California, San Diego

2006
To my parents, Qing Cao and Lihua Tang,

and my wife, Shuyan (Sunny) Sun.
# TABLE OF CONTENTS

Signature Page ........................................ iii
Dedication ........................................ iv
Table of Contents .................................... v
List of Figures ....................................... vii
List of Tables ....................................... viii
Acknowledgements ................................... ix
Vita, Publications, and Fields of Study .......... xi
Abstract ........................................... xii

Chapter I  Introduction  ................................ 1

Chapter II  Liquidity and Partial Liquidation in Debt Financing ........... 7
  II.A  Introduction .................................... 7
  II.B  Related Literature ............................... 12
     II.B.1  Empirical Evidence .......................... 12
     II.B.2  Dynamic Debt Contracting Models ............. 15
  II.C  The Model ...................................... 21
     II.C.1  The Setup .................................... 21
     II.C.2  The Optimality of PHSL Debt Contract ......... 27
  II.D  Discussion ..................................... 38
     II.D.1  Intuitions behind the PHSL Debt Contract ...... 38
     II.D.2  Implications .................................. 40
  II.E  Concluding Remarks ............................. 42

Chapter III  Financial Pressure and the Dynamics of Agency Costs ........ 45
  III.A  Introduction ................................... 45
  III.B  Empirical Framework ........................... 50
     III.B.1  Background .................................. 50
     III.B.2  Data and Variables ........................... 52
     III.B.3  Econometric Specification .................... 56
  III.C  Empirical Results .............................. 61
     III.C.1  Benchmark Case ............................... 62
     III.C.2  Robustness ................................... 66
     III.C.3  Tests of Alternative Hypotheses .............. 70
  III.D  Summary and Discussions ....................... 77
### III.E Appendix: GMM Estimation Results of the Benchmark Case  

### Chapter IV Asymptotic Distributions of Impulse Response Functions in Short Panel Vector Autoregressions  
- **IV.A Introduction**  
- **IV.B The Model and Estimation**  
- **IV.C Asymptotic Distributions of the IRFs**  
- **IV.D Simulation Evidence**  
- **IV.E Conclusion**  
- **IV.F Appendix: Proof of Theorem 10**

### Bibliography
<table>
<thead>
<tr>
<th>Figure</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>II.1</td>
<td>Time Diagram</td>
<td>25</td>
</tr>
<tr>
<td>III.1</td>
<td>The Impulse Response Functions to Shocks in FP1</td>
<td>63</td>
</tr>
<tr>
<td>III.2</td>
<td>The Impulse Response Functions in Alternative Cases</td>
<td>68</td>
</tr>
<tr>
<td>III.3</td>
<td>The Impulse Response Functions with Asset Sales in the VAR</td>
<td>71</td>
</tr>
<tr>
<td>III.4</td>
<td>The Impulse Response Functions with ARINV in the VAR</td>
<td>73</td>
</tr>
<tr>
<td>III.5</td>
<td>The Impulse Response Functions to Cash Flow and Interest Expenses</td>
<td>75</td>
</tr>
<tr>
<td>IV.1</td>
<td>The Impulse Response Functions of the First Variable to One SD Shocks</td>
<td>102</td>
</tr>
<tr>
<td>IV.2</td>
<td>The Empirical Coverages of Different Confidence Bands for N=100</td>
<td>104</td>
</tr>
<tr>
<td>IV.3</td>
<td>The Average Widths of Different Confidence Bands for N=100 and T=10</td>
<td>105</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table II.1</td>
<td>Acronyms for Different Contract Forms</td>
<td>26</td>
</tr>
<tr>
<td>Table II.2</td>
<td>Highest Level of Extractable Cash Flows</td>
<td>32</td>
</tr>
<tr>
<td>Table II.3</td>
<td>Gross Revenue to the Borrower</td>
<td>35</td>
</tr>
<tr>
<td>Table III.1</td>
<td>Summary Statistics of the Variables</td>
<td>57</td>
</tr>
<tr>
<td>Table III.2</td>
<td>Distribution of Firm Life</td>
<td>58</td>
</tr>
<tr>
<td>Table III.3</td>
<td>Analysis of Impulses — Benchmark Case</td>
<td>63</td>
</tr>
<tr>
<td>Table III.4</td>
<td>Analysis of Impulses — Alternative Case I</td>
<td>66</td>
</tr>
<tr>
<td>Table III.5</td>
<td>Analysis of Impulses — Alternative Case II</td>
<td>67</td>
</tr>
<tr>
<td>Table III.6</td>
<td>Analysis of Impulses — Alternative Case III</td>
<td>67</td>
</tr>
<tr>
<td>Table III.7</td>
<td>Arrelano-Bond (1991) System-of-Equations GMM Estimation Results</td>
<td>81</td>
</tr>
<tr>
<td>Table IV.1</td>
<td>AR Parameters for the Panel VAR(4) Model</td>
<td>100</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

To me, the most valuable part of achieving an accomplishment is not the final results but rather the experience of self-discovery along the journey. I owe my advisor, Garey Ramey, for this wonderful experience of making my original ideas into a rigorous model and testing it with real world data. His generous encouragement kept me patient and his luminous insights directed me out of dead ends. Garey also taught me many ways to improve my writing style and the organization of the paper materials. It is extremely lucky for me to have Garey as an advisor and it has always been great pleasure to work with him.

Professor Yixiao Sun spent many hours on coaching me the programming for the empirical chapter of this dissertation. He also carefully examined the manuscript of that chapter and gave many suggestions on how to make it more concise and readable. Chapter 3 of this dissertation is coauthored with Yixiao Sun in its entirety. All these facts clearly demonstrate the indispensability of Yixiao’s efforts in this work. The econometrics and the programming skills I learned from him are invaluable for this and all my future research projects. It is also fun to share stories about his young son. I am blessed to meet Yixiao as a teacher and a friend at UCSD.

The detailed questions and suggestions given by professor Michelle White benefit many parts of Chapters II and III in this dissertation. The discussions with professor Takeo Hoshi were always intellectually stimulating, warm and fun. Many conversations and emails with professor Bruce Lehmann broadened my views of financial economics, my research and my life and I highly appreciate his lovely attitudes towards almost everything. Professor Joel Watson was at one time on my committee and the contract theory I learned from him shapes the development of my model in Chapter II.
Many faculty members outside my committee also helped me a lot. Special thanks go to professor Valerie Ramey who taught me how to clean the raw data. The discussions from professors Roger Gordon, James Hamilton, Vince Crawford, Vincenzo Quadrini and Wouter Den Haan are sincerely acknowledged.

Many of my graduate classmates at UCSD provided much help in my life and research in La Jolla and they made the arduous tasks in a Ph.D. program bearable. Zhiwei Zhang, Liangjun Su and Zhigang Li helped me a lot in my early years at UCSD. Munir Andrés Jalil and Kazuki Onji are my best friends ever since my first year. Steve Scroggin and Philip Babcock sat next to me side by side in all the lectures for the first two years. Conversations with my officemates Munechika and Chung-Chiang were fun and relaxing. My undergraduate classmates Xiang Tang and Tao Wang also shared many their experiences in the study of economics and finance with me.

I would also like to thank my professors back to Beijing China, who encouraged me to pursue advance study in economics. They are professors Suiqi Lü, Jing He and Yi Liu from the School of Economics, Peking University, professor Qiren Zhou from CCER, Peking University and professor Angang Hu from Tsinghua University.

Finally, I want to thank my parents for their unconditional support, encouragement and love. My lovely girlfriend-wife Sunny Sun made my later years in the program a lot more colorful. The nice Chinese dishes she cooked had tremendous healing power to my home sickness. Many contributions made by Sunny are instrumental to the progresses in my career.
VITA

1976 Born, The City of Tianjin, People’s Republic of China

1999 Bachelor of Arts, International Economics and Trade
Peking University, Beijing, People’s Republic of China

1999–2000 Auditor, PricewaterhouseCoopers
Beijing, People’s Republic of China

2000–2005 Teaching Assistant, Department of Economics
University of California, San Diego

2003 Master of Arts, Economics
University of California, San Diego

2004–2005 Instructor, Department of Economics
University of California, San Diego

2006 Doctor of Philosophy, Economics
University of California, San Diego

PUBLICATIONS


FIELDS OF STUDY

Major Field: Economics

Studies in Microeconomics.
Professors Mark Machina, Ross Starr, Vincent Crawford and Joel Watson

Studies in Macroeconomics.
Professors James D. Hamilton, Marjorie Flavin, Valerie Ramey and Wouter Den Haan

Studies in Econometrics.
Professors Graham Elliot, James D. Hamilton, Halbert White, Clive Granger and Badi Baltagi

Studies in Financial Economics.
Professors Garey Ramey, Allan Timmermann, Bruce Lehmann and Ruth Williams
Chapter I introduces the research in this dissertation. In Chapter II, I build a model that links the partial liquidation pressed by debt obligations to the asymmetric information about the project quality and the agency costs of the borrower. Facing poor performance, the borrower in the model can sacrifice some future private benefit to generate cash by partial liquidation in order to fulfill his debt obligations. This action taken by the borrower can convince the lender of the long-term solvency of the firm. However, costly liquidations are inefficient, which is pertinent to the costs of business cycles during liquidity crunches.

Chapter III conducts a new test of the predictions of agency cost theory, based on the idea that episodes of financial pressure create dynamic incentives for managers to behave efficiently. This study adopts a panel-data vector autoregressions framework, using a large panel of firms, to estimate the dynamic responses of agency costs to financial pressure shocks. The paper shows that these dynamic responses are consistent with the intertemporal substitution effect in managerial agency costs implied by the model in
Chapter II. Tests on alternative hypotheses fail to reject the effects of financial pressure. The findings in Chapter III provide strong support to agency theories in the Corporate Finance literature.

Chapter IV establishes the limiting distributions of orthogonalized and nonorthogonalized impulse response functions in panel vector autoregressions with a fixed time dimension. The autoregressive parameters are estimated using the GMM estimators and the error variance is estimated using an extended analysis-of-variance type estimator. We find that the GMM estimator of the autoregressive coefficients depends on the estimator of error variance. This asymptotic dependence leads to additional terms in the asymptotic variance of the orthogonalized impulse response functions that are not present in the time series literature. Simulation results show that the asymptotic distribution of the orthogonalized impulse response function that takes the dependence into account is more accurate than the one that does not.
Chapter I

Introduction

Debt financing plays an important role in both the corporate finance literature and the real business cycle literature. In the corporate finance literature, Jensen and Meckling (1976) first relates agency costs to debt in the capital structure of a firm. A vast literature argues that using debt in the capital structure can restrain the agency costs in term of managerial perquisite consumption, although agency costs of debt also exist.\(^1\) Fisher (1933) explained the great depression with his debt-deflation theory. More sophisticated real business cycle models have been built to emphasize the effect of borrowing capacity variations on investments, which leads to persistent macroeconomic fluctuations.\(^2\) In this dissertation, I first establish a new dynamic debt financing model which provides new insights on the role of debt in both literatures. Then I examine the implications of this model using a large panel of firm level data and find consistent empirical evidence. To show the statistical significance of my results, new formulae in the

\(^1\)Examples for the former argument are Grossman and Hart (1980), Jensen (1986), Stulz (1990), Bolton and Scharfstein (1990), Hart and Moore (1995), Hart and Moore (1998) and Zwiebel (1996), among others. Examples for the latter argument are Jensen and Meckling (1976) and Myers (1977), among others.

\(^2\)For examples, Bernanke and Gertler (1989) and Kiyotaki and Moore (1997).
Panel data vector autoregressions are developed subsequently.

Chapter II introduces a divisible asset into the dynamic debt contracting literature. I assume that the borrower can convert this divisible asset into liquidity at some private cost. This setup gives the fulfillment of the interim repayment obligation an information role as to the long term solvency of the borrower’s project, especially when the short term performance of the project is poor. In the assumptions of the model, a good project holder can get high nonassignable future control rents by running the firm as a “going concern”. The repayment required by the debt contract in low-cash-flow states thus can force the holder of a bad project to default. This is because the nonassignable future control rent is not high enough to induce the borrower to sacrifice some private benefit now. In this way, the partial liquidation of the divisible asset also relates the intertemporal substitution of the managerial agency costs to the balloon risk of the debt, since I assume a bad project has zero value at maturity of the debt. On the other hand, a good project holder would like to sacrifice some private benefit now to generate some liquidity for fulfilling the interim repayment requirement in exchange for the high future control rent.

The private benefit that the good project holder has to give up for liquidity is a form of efficiency loss during liquidity crunches. This implication of the model links debt financing to the costs of recessions in business cycles. Since liquidity shocks are enough to generate inefficiency in this model, they can become an alternative driving force to the technology shocks in the real business cycles models. In general, the model also indicates that the financial flexibility of a firm can affect its borrowing capacity through the feasibility of costly signaling. Thus, the propagation effect of debt financing through the variation of borrowing capacity and investment levels can also be obtained by this model.
The costly signaling in Chapter II arises from the intertemporal trade-off in the managerial agency costs when liquidity is low. This is a dynamic restraining effect of debt whose empirical evidence has not been provided. In many models of managerial agency costs, debt can be used to restrain agency costs without complete ownership by management. The existing cross sectional tests on the effect of leverage on agency costs fail to provide convincing evidence to support or contradict this prediction. To address these two issues, Chapter III applies panel data vector autoregressions (VARs) to estimate the dynamics in managerial agency costs and adopts a direct liquidity measure as the proxy for financial pressure. The dynamics in managerial agency cost measures I find are consistent with the theoretical predictions.

The panel data VARs used in Chapter III can show exactly how managerial agency cost measures respond to unexpected liquidity shocks over time. Using the dynamic panel data estimator proposed by Arellano and Bond (1991), I can easily control for macroeconomic fluctuations and invariant firm characteristics that might generate spurious results. Under some structural assumptions as to the relationships between the variables in a vector autoregression system, I can extract the changes in managerial agency costs caused only by surprises in financial pressure levels. In the vector of variables, I include two agency cost measures, a proxy for financial pressure and a proxy for the growth opportunity of the firm. Thus the dynamic responses in the agency cost measures I find are due neither to the trends or shocks in the agency cost variables nor to simultaneous changes in both agency costs and financial pressure caused by financing behavior of a firm when exploiting its growth opportunities.

I use an interest-cash flow differential to measure financial pressure, instead of a direct measure of leverage. The reason is that high leverage does not necessarily imply high financial pressure and debt can generate its own agency costs. A more direct
measure of liquidity better captures the story described in Chapter II. The future control rent gives an incentive for the managers to reduce their current agency costs, thus firms with low future profit potential are excluded from the sample.

The managerial agency cost measures are the SG&A rate and asset turnover. The SG&A rate is ratio between the selling, general and administrative expenses and the beginning balance of the net fixed assets. Asset turnover is the ratio between sales and the beginning balance of the firm’s total assets. High SG&A rates and low asset turnover reflect high agency costs stemming from high perquisite consumptions and low effort levels of the management.

The dynamic responses in the empirical results show that after a surprise increase in financial pressure, the SG&A rate is significantly reduced below its long run average for two years but three years later it rebounds significantly above the long run average level. This dynamic behavior is consistent with the idea that managers reduce their current agency costs in order to protect their positions under financial pressure, since the future agency costs from these positions are high enough. The dynamic behavior of asset turnover after a positive surprise in financial pressure has a hump shape over time, i.e., asset turnover will rise above its long run average level for three years, starting from year two after the shock, before the impact of the financial pressure shock eventually dies out. This may reflect increased effort in boosting sales by the managers, or the reduction in assets used only for the purpose of managerial perquisite consumption, or both.\footnote{Some of my empirical findings are in favor of the first explanation for the dynamic behavior in asset turnover, i.e. higher sales increase this ratio.}

These results are robust in several alternative empirical formulations. Tests on alternative hypotheses show that the increase in asset turnover cannot be attributed to the asset sales under financial pressure. Although further test reveals the responses in
agency cost measures are largely due to the variations in cash flow, this conclusion does not dismiss the restraining role of debt financing, since debt appears to be the only known corporate governance device that provides hard constraints under poor performance. These findings support the broad implications of the model in Chapter II as well as other dynamic debt contracting models, such as Zwiebel (1996). The statistical significance of these results is assured by using the asymptotic distributions for the responses derived in Chapter IV.

The data set used in Chapter III contains a large panel of manufacturing firms from 1989 to 2003. The character of this panel is that its cross sectional dimension is large but its time series dimension is short. The estimation of a panel data VAR system on a short panel has broad empirical applications. Although the estimation of slope coefficients has been analyzed in the literature, the estimation of the impulse response functions (IRFs) and the analytical results for their confidence bands has yet to be provided. Professor Yixiao Sun and I derive these asymptotic confidence bands and examine their finite sample properties by simulation in Chapter IV.

We assume that the slope coefficients are the same across all individual units and there is no cross sectional dependence after controlling for the fixed individual and time effects. The autoregressive coefficients are estimated using the GMM estimators proposed by Anderson and Hsiao (1982) or Arellano and Bond (1991). Under our model specification, the equation-by-equation GMM estimators for both types are shown to be asymptotically equivalent to their corresponding system-of-equations estimators. The error variance is estimated by an extended analysis-of-variance type estimator. We show that the slope coefficient estimator and the covariance estimator are asymptotically de-

---

4 Examples include Love and Zicchino (2002), Gilchrist, Himmelberg, and Huberman (2005) and Chapter III of this dissertation among others.
pendent. This is due to the non-vanishing correlation between the demeaned regressors and the demeaned regression error when the time dimension is small.

We then show that the asymptotic variance of the orthogonalized impulse response functions contains additional terms because of the dependence between the two estimators. Simulation results show that the confidence bands based on the asymptotic distribution we derive are wider than alternative estimators that do not take into account this dependence. Thus, the confidence bands constructed without considering this dependence can find spurious statistical significance.

To summarize: the next chapter will model dynamic debt contracting with liquidation costs and agency problems; Chapter III tests the broad implications of that model on managerial agency costs under financial pressure; and Chapter IV provides the econometrics for panel data vector autoregressions that matters to the statistical significance of the results in Chapter III.
Chapter II

Liquidity and Partial Liquidation in Debt Financing

II.A Introduction

Modern treatments on Fisher (1933)’s debt-deflation theory have considered mechanisms through which debt financing transmits shocks in the economy via investment behavior constrained by borrowing capacity due to agency problems. Another important link in the “chain of consequences” in Fisher’s debt-deflation theory is the so called “distressed selling”. In this chapter, I model the forced asset sales under financial pressure caused by debt contracting. Many empirical studies on asset sales by firms under financial pressure find that this kind of sale is costly. In one of the most important modern model on liquidation, Shleifer and Vishny (1992) point out, most capital assets are specialized, and the assets sold during an industrial downturn may hardly be redeployed to their best usage. Misallocation thus results in low prices for the asset sales forced

1For example, Bernanke and Gertler (1989) and Kiyotaki and Moore (1997)
by liquidity needs. Researchers interpret their empirical findings as consistent with this argument. There is a noticeable gap between the facts in these empirical studies and the model used by Shleifer and Vishny (1992), however. Their model assumes complete liquidation of the firm in a liquidity crunch, yet most firms in these empirical studies only sold some of their assets rather than going through complete liquidation. Especially, Asquith, Gertner, and Scharfstein (1994) find that a higher portion of asset sales can help firms to avoid bankruptcy. Fisher (1933)’s theory also emphasizes the debt burden on the firms after “distressed selling,” so complete liquidation is not an assumption used by him. Then why do firms use costly partial liquidation during a liquidity crunch? My hypothesis is that this involves costly signaling. The simple debt contracting model I present, based on this hypothesis, can close the gap between theory and reality. My model can provide equilibria supporting an optimal debt contract that has the following features: i) the optimal debt contract requires a minimum fixed repayment, even when the borrower falls into a liquidity crunch; ii) this fixed repayment is linked to the balloon risk of the lender; iii) the borrower can avoid complete liquidation by liquidating a portion of the firm’s assets at the borrower’s own cost under this optimal debt contract; iv) furthermore, this optimal debt contract is renegotiation-proof. In addition, the “distressed selling” in my model directly generates inefficiency due to idiosyncratic income shocks. This implication enriches the role of debt in business cycles.

My modeling strategy is to embed a classical signaling game in the debt contracting problem. The borrower in my model can choose to liquidate a portion of the firm’s assets at his own costs in order to signal the long term solvency of his project to the lender.\(^2\) This helps to resolve the conflict of interests between the lender and the

\(^2\)I follow Aghion and Bolton (1992) throughout the paper refering to the borrower as he and the lender as she.
borrower when a negative liquidity shock hits the firm. If the lender views low cash flow in the short term as a sign of permanent decline in the firm’s future profitability, it may be her best interest to terminate the project now before the opportunity of partially recovering the loss disappears. But the borrower may want to continue the project for his private benefit no matter whether the firm is solvent or not in the future. With costly signalling, the borrower can credibly convey to the lender his private information about high future profitability of the firm. The lender then let the firm continue instead of terminating it. The borrower can benefit from cross-state subsidizing for doing this. The higher value recovered from liquidation of firms with low future profitability states allows the lender to demand lower repayments in other states. In this way, the total expected surplus to the borrower in financing the project with debt is maximized.

The costs in sending out the signals provide a new linkage between the use of debt financing and the welfare loss in recessions. Monetary policies trying to help businesses avoid liquidity crunches make sense when debt-forced partial liquidations cause efficiency loss. There are several sources of efficiency loss in partial liquidation. One popular form of efficiency loss is emphasized by Shleifer and Vishny (1992) in a general equilibrium outcome in a debt financing model analyzed in Hart (1993). An equilibrium can arise during an industry-wide downturn in that model: firms inside the industry don’t have enough cash to pay at their high valuation of specialized assets to that particular industry; but the outsiders with deep pockets just don’t have a high valuation on these specialized assets at all. The result is that the firm has to sell the assets to the outsiders who value their assets less. The asset seller has to accept a discount because of the inefficiency in misallocation. Direct evidence for this argument can be found in Pulvino (1998) and Ramey and Shapiro (2001), which I will discuss below.

Other possible sources of efficiency loss include the following. The market for
a firm’s assets may not be always liquid, and in order to sell the assets in a rush the seller may have to bear a pure liquidity discount\(^3\). Ramey and Shapiro (2001) also show that asset sales may involve searching and matching costs. Layoffs can be viewed as a particular type of liquidation, the liquidation of the firm’s specialized human capital. Shleifer and Vishny (1992) suggest laying off some workers in operations that are currently unprofitable may save some cash for fulfilling debt repayments. I call this the “cash saving layoff” hypothesis. The loss of firm specific human capital in the layoff, along with search and matching costs in labor markets, can cause efficiency loss in recessions too. In this way, the propagation of negative macroeconomic shocks can be explained partially as the ramification of the combination of the frictions in financial markets and labor markets. Ang, Cole, and Lin (2000), followed by Singh and Davidson (2003) suggest another reason for asset sales: under financial pressure, the managers may release firm assets that only serve for their perquisite consumption purposes. These kinds of asset sales actually improve the operational efficiency of the firm, but the sales may bear liquidity costs too. Moreover, as performance recovers, the managers may get their perks back by purchasing new assets at full price. When the old assets provided the same perks to the managers but they were sold at a discount and new ones are bought at full price, another example of potential efficiency loss emerges. But a startling feature in this story is that the managers actually sacrifice their own perk consumption (by definition, the perk consumption is a private benefit to the managers) to protect the firm from bankruptcy and thus protect their own positions. This is the exact assumption made in the model in this paper.

One problem with the analysis of Shleifer and Vishny (1992) is that they assume complete liquidation in their model. That is, firms in their model are filing for bank-

\(^3\)For example, see Kelly and LeRoy (2006).
ruptcy under Chapter 7. But the reality is that firms in financial distress more often sell part of their assets to raise cash for debt reduction rather than go directly to liquidation under Chapter 7. The findings in John, Lang, and Netter (1992), Ofek (1993) and Asquith, Gertner, and Scharfstein (1994), among others, reveal this fact. In John, Lang, and Netter (1992), only four of the 82 firms in their sample filed for Chapter 11 and only one liquidated but 29 firms sold assets or took similar actions. Ofek (1993) finds only four out of 358 financially distressed firms in his sample filed for bankruptcy in their first year of poor performance, but 82 firms sold assets or took similar actions. Asquith, Gertner, and Scharfstein (1994) study a sample of 76 junk bond issuers. In their sample, 21 firms sold more than 20 percent of their assets and 42 firms filed Chapter 11, but only one liquidation is mentioned.

I build a model that combines the partial liquidation phenomenon with the assumption that the costs in partial liquidation are born by the borrower. This model can show that partial liquidation not only provide liquidity to the borrower but also conveys information on the future solvency of the firm to the lender. This hypothesis is also supported by empirical evidence which will be discussed in detail in the next subsection.

To ensure that the minimum repayment requirement is not an empty threat, I make a pair of assumptions similar to the ones in Hart and Moore (1998). It is assumed that the lender has the right to foreclose the firm’s assets and obtain the complete liquidation value of them. On the other hand, the borrower can divert away the interim cash flow generated by the project. As a result, the interim repayment amount in the optimal debt contract needs to be renegotiation-proof. The lender can use the foreclosure right to enforce the interim payment but she cannot demand the repayment amount to be higher.

4Typical actions that are similar to asset sales are divestiture, spinoff, selling businesses of subsidiaries or discontinuing operations.
than the full liquidation value of the firm’s assets. A side benefit of these assumptions is that they make the model more tractable. To further simplify the analysis, I also assume that the state of cash flow at maturity is contractible, as in Hart (1993). Thus one may also interpret the costly partial liquidation as being enforced by debt overhang. I further assume the firm’s assets have no liquidation value at maturity, and if the project quality is bad, there will be no revenue accruing to the borrower. The lender then faces the possibility of losing a significant repayment amount at the maturity of the loan, i.e. the balloon risk in debt. Separating by signaling in this model then adds value because in a certain state, the project can be terminated at liquidation value before the realization of a huge loss at maturity.

The remainder of this chapter is organized as follows: Section II.B discusses empirical evidence that motivates the model in this paper and relates it to the literature on dynamic debt contracting. Section II.C first describes the setup of the model, then analyzes the model and gives the set of conditions that can support the optimal debt contract as described. Section II.D discusses the intuitions and implications of the model. Section II.E concludes the paper.

II.B Related Literature

II.B.1 Empirical Evidence

The inefficiency in asset sales under financial pressure is supported with both direct evidence and reflected in the stock market reactions to the announcements of asset sales. Pulvino (1998) studies a large sample of used aircraft transactions occurring from 1978 to 1991. The study shows that when the airline industry is depressed, airlines
with low spare debt capacities sell aircraft at a 14 percent discount to the average market price. Also, during market recessions, the sellers with tighter financial constraints are more likely to sell their used aircraft to financial institutions. As industry outsiders, the financial institutions pay a discount of 30 percent to the average market price during market recessions. Evidence on the pattern of airlines’ used aircraft purchases also supports the hypothesis that financially constrained airlines liquidate aircraft at discounts to fundamental values.

Ramey and Shapiro (2001) study aerospace plant closings with equipment-level data. They estimate that the average market value of equipment for the sample is 28 cents per dollar of replacement cost. Types of capital that they identify as being more specialized sell at a greater discount. The capital that sold to industry insiders sells at significant premia, as large as 100 percent, compared to that sold to outsiders. This supports Shleifer and Vishny’s view that the industry insiders value the capital specialties more than the outsiders, too. Moreover, Ramey and Shapiro (2001) find that in private liquidation sales with costly search and matching, the assets are sold at a substantial premium over those at the public liquidation. Thus it takes time and is costly to sell the assets to a buyer who actually values more the capital specificities especially in a thin and illiquid market.

Studies of stock price reactions to the announcement of asset sales also provide evidence consistent with the liquidation costs identified by Shleifer and Vishny (1992). Brown, James, and Mooradian (1994) find that if the proceeds of the asset sales are announced to be debt repayments, the abnormal stock return is significantly lower, by 3.17%. Moreover, the low abnormal return is even more statistically significant if the firm successfully avoided bankruptcy with the help of the sales. This means that even the good news of survival of the firm cannot compensate the shareholders for their value
loss. The value loss in the inefficient liquidation forced by creditors will be mostly born by the shareholders. Lang, Poulsen, and Stulz (1995) find contrary results in stock reactions to asset sales. The firms in their sample are healthier firms than those in the sample of Brown, James, and Mooradian (1994). Lang et al (1995) interpret their findings as consistent with Shleifer and Vishny’s arguments too, because according to Shleifer and Vishny’s story, good deals in selling assets are hard to find. Successful asset sales are good news to shareholders and thus have a positive impact on stock returns. Asquith, Gertner, and Scharfstein (1994) also provide some indirect evidence for Shleifer and Vishny’s idea. Their sample is composed of financially distressed junk bond issuers. Asquith et al. find that companies in poorly performing or highly leveraged industries are less likely to sell assets, so the good deals are indeed hard to find. Similar to long term assets, deep discount on short term assets, such as inventories and accounts receivable may be accepted by a distressed firm due to lower valuation or just bargaining power from buyers.\(^5\)

Although firms do sell assets under financial pressure, they mostly engage in only partial liquidation, rather than immediately go through complete liquidation under Chapter 7, when negative liquidity shocks hit. As mentioned at the beginning, Asquith, Gertner, and Scharfstein (1994) find that asset sales can actually help firms avoid filing for bankruptcy. Only 14 percent of the firms that sell a large portion (more than 20%) of their assets go to bankruptcy, compared with 49 percent of the firms with small or no asset sales. For the sample of airlines in Pulvino (1998), I provide some updates based on information available on the websites of the airlines. Among the 14 airlines who sold more aircraft than they purchased in the sample of Pulvino (1998), six of them

\(^5\)I thank professor Ben J. Sopranzetti at Rutgers Business School – Newark and New Brunswick for substantiating this claim.
are still in business as of the year 2004, four merged with or were acquired by other airlines, and the remaining four airline companies were liquidated. In the sample of Brown, James, and Mooradian (1994), some firms that sell assets can successfully avoid bankruptcy. Thus in reality, complete liquidations are not necessary for many firms as described in Shleifer and Vishny (1992) though they still engage in costly asset sales. The implication of my model and Shleifer and Vishny’s general equilibrium story in aggregate liquidity crunches gives “distressed selling” an interpretation in line with the prisoners’ dilemma: Every firm with debt will try to survive by selling some assets. But all firms put their assets on the market at the same time, further attenuating liquidity in the economy.

II.B.2 Dynamic Debt Contracting Models

The model in this paper belongs to the literature on optimal dynamic debt contracting. There are two important features of this model. First, it is costly to provide information via partial liquidation to meet the debt repayment requirement. The costs directly relate debt financing to efficiency loss. Second, the information that is revealed pertains to the balloon risk to the creditors. Comparing to the existing literature, this model demonstrates these two features explicitly.

The efficiency issue is usually linked to the project quality chosen by the borrowers when financing with debt, as in Hart and Moore (1995) and von Thadden (1995).

---

6American Airlines, United Airlines, Northwest Airlines, Delta Airlines, Aloha Airlines and Hawaiian Airlines.
7Western Airlines, Trans World Airlines, Pacific Southwest and AIRCAL.
8Air Florida, Eastern Airelines, Pan Am World Airways and Braniff Airlines.
9Of course these facts may reflect that large firms are less likely to file for bankruptcy under Chapter 7, see White (1994) and it is also rare for them to liquidate after they file Chapter 11 bankruptcy, see Wruck (1990) The firms in these samples are indeed large firms as their financial information are available to the researchers.
Hart and Moore (1995) consider the role of long-term debt in preventing management from financing negative NPV “perk” projects. The indivisibility of the final revenue from the firm’s portfolio of projects gives long-term debt a role in constraining the investment choices made by the managers. Hart and Moore (1995) ties the project quality to the final repayment, whereas the model of this chapter ties the project quality to the interim repayment. In equilibrium, there is no inefficient investment in Hart and Moore (1995) because the debt obligation at time 2 can induce the managers to make efficient choices. However, in my model, as will be shown, the enforced interim repayment can result in efficiency loss despite the fact that fulfilling it can help the firm to survive.

von Thadden (1995) considers about the possibility that outside financing may induce inefficient myopic investment. He studies the debt financing for a project with a two-period investment requirement and two periods of returns. The borrower can privately choose the duration of the project, either long- or short-term. The short-term project has higher short term expected returns, but lower overall expected returns. Since short term poor performance cannot fully reveal the project quality, using the liquidation threat as a screening device will eliminate some good projects, which have high long term expected return. Monitoring by investors in this case can help to overcome this kind of short-term bias of investment. But my model emphasizes the unavoidable efficiency loss in the partial liquidation costs in the signaling game. This provides a stronger linkage between debt financing and efficiency loss in macroeconomic fluctuations. This inefficiency effect also arises under less restrictive conditions than needed in Shleifer and Vishny (1992). Idiosyncratic revenue shocks in this model are enough to generate efficiency loss, whereas industry-wide depression is needed in Shleifer and Vishny (1992). This is because the potential sources of liquidation costs are more numerous than those considered by Shleifer and Vishny (1992), as pointed out in the Introduction.
of this chapter.

The model in this paper relates the steady repayment stream in debt financing to the long-term solvency of a debt. In the terminology of debt financing, my model states that the timely fulfillment of coupon payments or sinking fund provision by the borrower gives some confidence to the lenders that the principal will be paid back at maturity. Some of the previous literature did not adequately address the relationship between fixed payments and long-term solvency. For example, Chang (1990) refers to the steady repayment stream as a regular test of the firm’s financial strength. The dynamic financing model he built can generate a debt-looking contract consistent with this regular test idea. However, the money that can be saved by an early testing is only the verification cost in the later period. This will be the case only when the verification cost is increasing in the final period value of the firm’s assets. Indeed, in his model, the earlier payments are not related to, and thus say nothing about, the prospects for future repayments from the borrower’s project. The dynamic debt financing models in Hart and Moore (1994) and Hart and Moore (1998) assume there is no repayment obligation by the end of the project, as the borrower can always repudiate or divert the cash flow away at that time. In this sense, these two models cannot directly address the balloon risk without further interpretation.

For costly partial liquidation to provide balloon risk information, my model provides one important insight on debt modeling strategy: the fluctuation in cash flow cannot dependent solely on the project quality. Otherwise, the bankruptcy decision can just be contingent on the level of short term cash flows, rather than contingent on whether the fixed payment requirements are fulfilled or not. In other words, the actual payments of interests or sinking funds would not be important and credible cash flow levels are good enough in a debt contract. This makes my model different from the model of
incomplete contracting based on a noisy signal studied by Aghion and Bolton (1992). Their model also has the two period timing, but there is a short-term signal imperfectly correlated with the state of nature. Thus, a contract contingent on this noisy signal is incomplete. Because of the private benefit to the project manager, the allocation of project control rights matters for maximization of the total return. When contingent allocation of control rights is optimal, the financing contract in Aghion and Bolton (1992) has some similarity to a standard debt contract. In contrast, the shift of control rights, i.e., bankruptcy, in the model of this paper, is rather a cost-benefit consideration by the borrower. Private information is revealed by the choices made by the borrower rather than inferred from some signal as in Aghion and Bolton (1992) or cash flow level as in von Thadden (1995).

There are several similar assumptions and implications between my model and the models studies in Bolton and Scharfstein (1990), Diamond (1991), Hart and Moore (1994) and Hart and Moore (1998). Thus it is worth looking at the distinctions between this model and these other dynamic debt contracting models.

Liquidation threats play an important role in both the model in this paper and the one given by Bolton and Scharfstein (1990). Bolton and Scharfstein (1990) use liquidation threats to induce debt repayment. The model of Bolton and Scharfstein has three dates, with financing required at both times 0 and 1. Realized cash flows in each period are random, and the borrower can misreport the amount of cash received. Thus, only the lowest cash flow realization can be contracted on. Bolton and Scharfstein give the lenders all the bargaining power in their model. The threat to deny second-period funding may induce truthful reporting of a high first-period cash flow, allowing the lender to extract more cash at time 1. Comparing to my model, since there is no asymmetric information on project quality in their model, the repayment enforced by the liquida-
tion threat does not yield any information about the risk associated with the balloon payment. Furthermore, Bolton and Scharfstein’s model appears more like separated short-term debts rather than a long-term debt with fixed repayments and balloon risk, as in this paper.

The private control rent is explicitly modeled in this paper so as to give incentives to the borrower to keep the firm as a “going concern.” This is different from Aghion and Bolton (1992) and Hart (1993), where the private control benefit is simply assumed. But the control rent assumed here works in the same way as in Diamond (1991): it is accrued privately to the manager and is only awarded when the project runs to the second period. Cross-type subsidy also appears in Diamond (1991). In his model, the borrowers with good projects can use maturity choice to lower their expected financing cost by extracting value in the liquidation of borrowers with bad projects. But Diamond (1991)’s focus is to study the maturity structure choice of the firms with private information on project quality. In his setup, asymmetric information exists at time 0, and the lenders can receive an information update, or “credit rating,” on the borrower’s type at time 1. This credit rating affects refinancing costs differently for different maturity structure choices. The maturity structure that gives lowest overall financing costs for the good-type borrower will be chosen by both types of the borrower, since otherwise the bad-type borrower will reveal his type to the lender and thus cannot raise funds at time 0. This means only pooling equilibria exists in the Diamond model, and it does not allow for signaling at time 1. The interim repayments are omitted in his model as well in the long term debt. But my model analyzes the interim repayments in a debt contract, and these repayments are assigned an informative role.

Both of the debt contracting models in Hart and Moore (1994) and Hart and Moore (1998) have all repayments determined by renegotiation, as does the model in this
paper. But the repayments in Hart and Moore (1994) have a different meaning. In Hart and Moore (1994), the borrower can threaten to walk away from the project at any time. The amount the borrower can borrow may then be constrained by the liquidation value obtainable by the lender. Thus the interim payments are made for the purpose of keeping the outstanding liability within the limit of repudiation proof borrowing capacity. There is no uncertainty in Hart and Moore (1994), and interim default never actually occurs in the renegotiation-proof equilibria of their model. On the other hand, my model shows that random cash flow and liquidation of bad projects imply that interim default can happen under the optimal debt contract. Hart and Moore (1998) allow for random cash flow in both periods, along with an interim liquidation option, and there is again no asymmetric information about project quality. In contrast to the model in this paper, Hart and Moore permit the borrower to take out a loan with initial size larger than the actual required investment, as a cushion against “rainy days.” The borrower faces a higher payback obligation for this cushion when the project return is high. However, the higher payback obligation itself has no cost or efficiency implications. In this model, additional liquidity from outside creditors to the borrower can also blur the information content of cash repayments, and thus the current model implies a negative covenant on the borrower’s ability to borrow more debt.\footnote{See Fabozzi (2000) Chapter 9 for detailed discussion on debt covenants.}
II.C The Model

II.C.1 The Setup

I assume an entrepreneur, or the borrower hereafter, has exclusive access to a project and has no initial personal wealth. To set up the project, the borrower must raise funds from the lender. Both the lender and the borrower are assumed to be risk neutral. The lender assumes limited liability on his loan. A risk free return rate is available to both parties. Following the literature, this rate is normalized to zero. The financing contract takes the form of a debt contract in the sense that if default on the repayment occurs, the lender has the right to liquidate the project. To set up the project, the borrower uses the loan to purchase an indivisible asset. Thus the borrower must pay for it in full, and this asset can only be sold as a whole in case of liquidation. After the project is financed, the quality of the project is determined, which may be either good or bad. The project quality is private information to the borrower and it cannot be verified by outsiders, such as a court.

The project lasts for two periods. In other words, all actions of both parties will happen on three dates and denoted as dates 0, 1, and 2. Cash flows accrue to the borrower at the end of each period. The first period cash flow is random and it can be either high or low. I assume the state of the first period cash flow is verifiable, i.e. the first repayment can be based on the states of the cash flow. But this cash flow is completely divertible by the borrower, i.e., the borrower can divert funds to personal use on a one-for-one basis. The borrower also need not reserve any portion of the first

\footnote{Additional personal wealth from the borrower of course will have the effect of enhancing initial borrowing capacity and the liquid asset used as cushion may also be financed from the beginning, however the first problem has been studied by Gale and Hellwig (1985) and the second by Hart and Moore (1998) To make the model simpler, we set it up without involving these two problems.}
period cash flow to make contracted payments on date 2. If the borrower remains with the project into the second period, on date 2, a second contractible cash flow may be realized, and the borrower will also obtain a nontransferable control rent. The amounts of both the second period cash flow and the control rent depend on the project type.

To capture the partial liquidation of firm assets, I assume there is a divisible asset endowed to the borrower once the project starts. This assumption can be viewed as the accumulation of the firm’s specialized capital. Some portion up to the full amount of this asset can be liquidated at the end of the first period at the cost of the borrower. If the any portion of the divisible asset survives to date 2, the remaining portion can realize its full value at no cost. The borrower can divert all the proceeds from liquidating this asset on either date 1 and date 2, thus the cash flow generated this way is not contractible.

If the borrower defaults on the first period repayment, the lender will terminate the project and both the indivisible and divisible assets have to be liquidated. The indivisible project asset has only its liquidation value, which is lower than the initial investment. The lender can seize this liquidation value from the indivisible asset. The cost in liquidating the divisible asset is permanent if this happens on date 1 and it can never be recovered. The debt contract gives the lender the right to liquidate the project as a threat to enforce the first period repayment. As a result, the borrower may use the proceeds from liquidating the divisible asset to fulfill some repayment obligation on date 1 if he has enough incentive to do so. However, on date 2, there is no means for the lender to enforce the borrower to pay out his divertible cash flow. I also allow renegotiation of the contracted repayments, and all the bargaining power is allocated to the borrower. Based on these assumptions and the properties of cash flows from different sources on different dates, the following observations on the repayments to the lender are made:

I. The sources of the first period repayment can be from the cash flow to the
project and the proceeds from liquidating the divisible assets. The first repayment is enforced by the termination threat from the lender and the repayment amount is subject to renegotiation.

II. Since the second period cash flow is contractible and all other benefits that accrue to the borrower are not, the second repayment only depends on the contracted state-contingent amount out of the realized second period cash flow.

In designing the debt contract, we only need to satisfy the breakeven constraint for the lender, as usual in the literature. The borrower will choose a debt contract that maximizes his ex ante net payoff, subject to the lender’s breakeven constraint. Since the expected payout to the lender is fixed, the borrower thus will simply choose the debt contract that generates the highest level of gross benefits from all sources. In this sense, the borrower always prefers to get the project funded since he has nothing to lose in this game. This reflects the borrower’s limited liability constraint.

To begin our analysis, I introduce the symbols into the story above. On date 0, if the lender and the borrower sign the debt contract and the loan amount of $K$ is transferred for purchasing the indivisible asset, then the project is set up.

On date 1, four things happen:

1. The quality of the project is realized. The quality can be good, denoted as $G$, with probability $\rho$, or bad, denoted as $B$, with probability $b = 1 - \rho$ and $1 > b, \rho > 0$. The project quality is private information to the borrower.

2. The project yields a short-term random nonnegative cash flow, received by the borrower. When the amount is positive, we denote it as $X_1$. We denote this state as $H$ and it happens with probability $h$, where $0 < h < 1$. $X_1$ is fixed and known to both parties. Yet, with probability $l = 1 - h$, the date 1 cash flow can be zero. We denote this state as $L$. This fluctuation can happen to both types ($G$ or $B$). Therefore, we will
have four states – GH, GL, BH, BL – with nonnegative probability to occur.

3. The indivisible project asset can be liquidated at time 1 at a value $\alpha K$, with $0 < \alpha < 1$. This is a decision to be made by the lender if the borrower defaults on date 1. The indivisible asset is assumed to be worthless at time 2.

4. The borrower will be endowed with a divisible asset $D$. Since the asset is divisible, we use $\gamma$ to denote the liquidated portion, with $\gamma^H$ for high cash flow state, $\gamma^L$ for low state and $0 \leq \gamma^H, \gamma^L \leq 1$. The cost or inefficiency of this liquidation lies in the price $\beta$ at which $D$ can be sold on date 1 and we assume $0 < \beta < 1$. On date 2, $(1 - \gamma)D$, the portion remaining from last period, can be sold at unit price 1. Thus $1 > \beta$ reflects the cost of premature liquidation on date 1.

The cash flow generated by the project at date 2 depends on its quality. If the project is of type G, the cash flow will be $X_2 > 0$; but if the project is of type B, time 2 cash flow will be zero. I assume the borrower obtains a nonassignable control rent if the project runs through the second period. This means that the borrower can obtain the control rent on date 2 only if the project is not liquidated on date 1. The value of this control rent will be $C$ for state GH, GL and BH. This assumption on $C$ means the borrower can enjoy control rents from good projects. If the project fails eventually, as long as his track record is good, he can find enough excuses to deny his responsibility for the failure. When the state is BL, i.e., a low first period cash flow to a bad quality project, this control rent will be $v$. I assume $v < C$, reflecting that a bad track record and final business failure will reduce the control rent to the borrower.

By our assumption, the debt contract cannot be written on the project quality, but the state of cash flow on date 1 is contractible. The contract specifies four repayments contingent on the cash flow state on date 1, $R_1^H$ and $R_1^L$ on date 1 and $R_2^H$ and $R_2^L$ on date 2. All $R_1$’s and $R_2$’s are nonnegative. Notice that the $R_2$’s are only for type G
projects as type B projects has no cash flow to divide on date 2. While the $R_2$’s are directly enforceable, the $R_1$’s can be enforced only through the threat of liquidation on date 1, following Hart and Moore (1998). Since both the lender and the borrower are risk neutral and the borrower has all the bargaining power, as long as the project is funded, the borrower can reap all the ex ante surplus above the investment $K$, i.e., present value of the expected cost for the project. Once the project is funded, this surplus depends only on the parameter values, regardless of $R_1$’s and $R_2$’s specified in the contract. The time diagram is shown in Figure II.1.

For a certain subset of the parameter space, the $R_1$’s can take values that force the borrower to default in some states. In this way, the borrower can separate himself from bad types. On date 1, by the bargaining power of the borrower and renegotiation, the highest enforceable amount for the $R_1$’s is $\alpha K$. Obviously, terminating all types of project on date 1 or letting only type B projects continue does not make sense, since the expected payout to the lender must be no more than $\alpha K$. However, all the contract

---

12Recall that for type B, there will be zero second period cash flow thus no repayment on date 2 to the lender. Thus if a debt contract force all types to terminate on date 1 or only type B projects continue, the repayment to the lender only happens on date 1 in all states at the amount no more than $\alpha K$. 
forms in the Table II.1 have to be considered for an arbitrary set of parameter values. This is because although some of the contract forms are seemingly dominated by others (for example, GLO compared to GHO), they may become optimal in some nontrivial parameter space. To facilitate the analysis, all the possible contract forms are given acronyms in Table II.1.

<table>
<thead>
<tr>
<th>Acronyms</th>
<th>Continuation Types</th>
</tr>
</thead>
<tbody>
<tr>
<td>States</td>
<td>GH</td>
</tr>
<tr>
<td>GHO</td>
<td>Yes</td>
</tr>
<tr>
<td>GLO</td>
<td>No</td>
</tr>
<tr>
<td>SHSL</td>
<td>Yes</td>
</tr>
<tr>
<td>PHS L</td>
<td>Yes</td>
</tr>
<tr>
<td>SHPL</td>
<td>Yes</td>
</tr>
<tr>
<td>PHPL</td>
<td>Yes</td>
</tr>
<tr>
<td>PH</td>
<td>Yes</td>
</tr>
<tr>
<td>PL</td>
<td>No</td>
</tr>
</tbody>
</table>

Given the game above, our debt contracting problem becomes choosing $R_1$’s and $R_2$’s to maximize the borrower’s expected return subject to incentive compatibility constraints for the borrower under different states and the breakeven constraint for the lender. In addition, the contracted payments must satisfy feasibility and renegotiation-proofness constraints. Based on the assumptions in the setup described above, we can further observe the following properties of the optimal debt contract.

**Lemma 1** The optimal debt contract has the following properties:

a) every debt contract form has to minimize the inefficient liquidation of the divisible asset;

b) the borrower will choose the debt contract form that maximizes the expected gross revenue to the project.
Due to the high dimension of the problem, the next subsection only provides a set of sufficient conditions that makes the most interesting debt contract form optimal.\textsuperscript{13} The debt contract form we are interested in is the one that obtains pooling outcomes in state H and separating outcomes in state L (PHSL). Usually, as long as the borrower has no trouble fulfilling the interim repayment obligation such as in state H, the lender will not take any action. When first period revenue is zero (in L state), the borrower may still want to hold on to the project if he knows the project is of type G; but the lender may not know that and thus want to force him to liquidate the project. When it is possible for the borrower to use costly liquidation to separate himself in this case, the liquidation of the type G project may be avoided even the first period cash flow is zero. This case has important macroeconomic implications since the costly liquidation in response to negative shocks may propagate inefficiency into the whole economic system.

\section*{II.C.2 The Optimality of PHSL Debt Contract}

To focus our attention to the PHSL debt contract, I further assume:

\[ \min\{X_1, C, X_2\} > \alpha K. \] \hspace{1cm} (A1)

Given that \( X_1 > \alpha K \), by renegotiation-proofness, the first repayment in the H state can be at most \( \alpha K \) and no costly liquidation will be involved. Formally, the renegotiation-proof condition can be stated as:

\[ R^H_1, R^L_1 \leq \alpha K. \] \hspace{1cm} (II.1)

A first consequence of the above two formulae is the following:

\textsuperscript{13}The analyses of simpler cases are available upon request. These simpler cases include the models without asymmetric information, the divisible assets or the first period income uncertainty.
Lemma 2  Given assumption A1 and condition (II.1), in state H, there will be no liquidation of the divisible asset D.

Proof. : Because the borrower has enough cash in state H to pay the highest amount that can be requested from the lender, by assumption A1 and condition (II.1), lemma (1) ensures this conclusion. ■

Also since C > αK, there is enough private benefit to attract GH, GL and BH types to stay with the project until date 2. This is because the control benefit for the borrower, C, is higher than the highest amount he has to pay out on date 1, αK. Then, a second consequence from the assumption (A1) is that separation in state H is no longer possible, thus debt contract forms GHO, SHSL, SHPL are out of consideration. Thus, the only contract forms GLO, PHSL, PHPL PH and PL are left for consideration because termination of the project in state L is still possible to enforce.

Moreover, since it is unnecessary for the lender always to let both the types (GH or GL) continue, \(X_2 > \alpha K\) further simplifies our analysis. \(X_2 > \alpha K\) allows the lender to be persuaded not to terminate the project as long as the borrower can convince her the project is of type G. This gives the reason why the debt contract may allow the project in state GL to run through the second period. As we will show later, this intuition is reflected in the fact that \(X_2 > \alpha K\) assures that the PHSL contract has the highest borrowing capacity.

A debt contract that can support the PHSL outcome has to satisfy the following conditions.

\[
X_1 + \beta D \leq X_1 + X_2 + C + D - R^H_1 - R^H_2 \quad \text{(II.2)}
\]

\[
\beta D \leq \beta \gamma^L D + X_2 + C + (1 - \gamma^L)D - R^L_1 - R^L_2 \quad \text{(II.3)}
\]

\[
X_1 + \beta D \leq X_1 + C + D - R^H_1 \quad \text{(II.4)}
\]
\[ \beta D \geq \beta \gamma^L D + v + (1 - \gamma^L)D - R^L_1 \quad \text{(II.5)} \]

\[ \rho h(R^H_1 + R^H_2) + \rho(1 - h)(R^L_1 + R^L_2) + (1 - \rho)h R^H_1 + (1 - \rho)(1 - h)\alpha K = K \quad \text{(II.6)} \]

\[ R^H_1 \leq \alpha K \quad \text{(II.7)} \]

\[ R^L_1 \leq \min\{\beta D, \alpha K\} \quad \text{(II.8)} \]

\[ R^H_2, R^L_2 \leq X_2 \quad \text{(II.9)} \]

In an optimal PHSL debt contract, \( R^H_1, R^L_1, \gamma^L, R^L_2 \) and \( R^H_2 \) will be chosen to maximize the expected surplus to the borrower. Conditions (II.2) to (II.4) are the incentive compatibility conditions for the borrower in states GH, GL and BH to continue. These conditions guarantee that the benefits from fulfilling the contracted repayments are higher than benefits from default. Condition (II.5) is also an incentive compatibility condition. But it is the condition that forces the default of type B project holder in state L. Condition (II.6) is the breakeven condition for the lender. Conditions (II.7) and (II.8) reflect the renegotiation-proofness of the first repayment amount. In addition, condition (II.8) also reflects the feasibility condition for the borrower to make the payment with available cash. Finally, condition (II.9) is the feasibility condition for the second repayment amount. These feasibility conditions result from the limited liability assumption for the borrower.

We can simplify conditions (II.2) — (II.5) to take advantage of assumption (A1) for our analysis:

\[ R^H_1 + R^H_2 \leq X_2 + C + \delta D \quad \text{(II.10)} \]

\[ R^L_1 + R^L_2 \leq X_2 + C + (1 - \gamma^L)\delta D \quad \text{(II.11)} \]

\[ R^H_1 \leq C + \delta D \quad \text{(II.12)} \]

\[ R^L_1 \geq v + (1 - \gamma^L)\delta D \quad \text{(II.13)} \]
where $\delta = 1 - \beta > 0$. By our assumptions $C > \alpha K$, $D > 0$, conditions (II.7) — (II.9) and $\gamma^L \in [0, 1]$, conditions (II.10) — (II.12) are obviously satisfied, as we have argued before.

Conditions (II.7) — (II.9) only depend on the parameter values and together they will determine the borrowing capacity for the borrower through (II.6). Condition (II.13) can be transformed to:

$$R_1^L = \beta\gamma^L D \geq \beta(v + \delta D).$$

(II.14)

The first equality is a condition for the optimal debt contract imposed by lemma 1, which simply means that there is no need to incur the inefficient liquidation more than necessary to fulfill the first repayment requirement in state L. Condition (II.14) then implies

$$\beta D \geq v,$$

(II.15)

following the fact that $0 \leq \gamma^L \leq 1$. Also, the renegotiation proof condition (II.8) and condition (II.14) demands

$$\alpha K \geq \beta(v + \delta D).$$

(II.16)

For the rest of our analysis, I will assume both (II.15) and (II.16). Also by lemma (1), if strict inequality needs to hold in condition (II.14) for purposes of satisfying the breakeven condition (II.6), then condition (II.9) must hold at equality since otherwise, one can always increase $R_2$’s to lower the inefficient liquidation of the divisible asset on date 1. When condition (II.9) holds at equality, condition (II.11) implies

$$R_1^L = \beta\gamma^L D \leq \beta(C + \delta D).$$

(II.17)

Combining (II.14) and (II.17), we have $\beta(v + \delta D) \leq R_1^L \leq \beta(C + \delta D)$. It is helpful to mention the intuition behind $\beta(v + \delta D)$ and $\beta(C + \delta D)$. These two formulae are nothing
but the private benefits to the borrower for running the project until date 2, discounted at price $\beta$. The meanings of $v$ and $C$ are obvious and $\delta D = (1 - \beta)D$ reflects that the borrower can obtain $\beta D$ for sure by default on date 1, so only the premium from liquidating $D$ at maturity matters. To further simplify the situation, I assume the value of complete liquidation of the divisible asset on date one is less than the liquidation value of the indivisible asset, i.e.,

$$aK \geq \beta D.$$  
(II.18)

By assumption A1 and condition (II.15), we then have the following ranking among the parameters:

$$aK \geq \beta D \geq v.$$  
(A2)

I make A2 another assumption for the rest of my analysis since it combines conditions (II.15) and (II.18) and it implies condition (II.16) and condition (II.17) when combined with A1. Assumption A2, conditions (II.14) and (II.15) imply condition (II.16). For condition (II.17), assumption A1 and A2 imply $C > aK \geq \beta D$ and $C > \beta D$ implies $\beta(C + \delta D) > \beta D \geq \gamma^L \beta D$.\textsuperscript{14} I summarize the results on the first repayment in state L in the following lemma:

**Lemma 3** Given assumption A1 and A2, in an optimal PHSL debt contract, $R^L_1 \in [\beta(v + \delta D), \beta D]$.

To compare the PHSL debt contract with other debt contract forms, we need to specify the conditions that let the type B project holder pool with the type G project holder in state L:

$$R^L_1 = \beta \gamma^L D < \beta(v + \delta D).$$  
(II.19)

\textsuperscript{14}For $\beta(C + \delta D) > \beta D \iff \beta C + \beta \delta D > \beta D \iff \beta C + \beta D - \beta^2 D > \beta D \iff \beta C > \beta^2 D \iff C > \beta D.$
In this condition, I assume a type B project holder will only continue when the first repayment in state L is strictly less than the private benefits to himself. For the debt contract forms that admit pooling outcomes in state L, we have the following lemma for the first repayment in state L:

**Lemma 4** Given assumption A2 and condition (II.1), in a debt contract allowing type B borrowers to pool with the type G borrowers, \( R^L_1 \in [0, \beta(v + \delta D)) \).

To check whether a debt contract will be signed by the borrower and the lender, the borrowing capacity from the debt contract has to be able to support the initial outlay, \( K \). At the same time, the optimal contract has to satisfy lemma 1. The borrowing capacity of a debt contract form is determined by the highest level of extractable cash flow from the borrower. Using Table II.2 makes it very convenient to calculate and compare the borrowing capacity for different contract forms.

<table>
<thead>
<tr>
<th>Acronym/Date</th>
<th>States [prob.]</th>
<th>GH [( \rho h )]</th>
<th>GL [( \rho l )]</th>
<th>BH [( bh )]</th>
<th>BL [( bl )]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>States [prob.]</td>
<td>1  2</td>
<td>1  2</td>
<td>1  2</td>
<td>1  2</td>
</tr>
<tr>
<td>PHS1L</td>
<td>( \alpha K )</td>
<td>( X_2 )</td>
<td>( \beta D )</td>
<td>( X_2 )</td>
<td>( \alpha K )</td>
</tr>
<tr>
<td>PHPL</td>
<td>( \alpha K )</td>
<td>( X_2 )</td>
<td>( \beta(v + \delta D)^- )</td>
<td>( X_2 )</td>
<td>( \alpha K )</td>
</tr>
<tr>
<td>GLO</td>
<td>( \alpha K )</td>
<td>0  ( \beta(v + \delta D)^- )</td>
<td>( X_2 )</td>
<td>( \alpha K )</td>
<td>0  ( \alpha K )</td>
</tr>
<tr>
<td>PHS</td>
<td>( \alpha K )</td>
<td>( X_2 )</td>
<td>( \alpha K )</td>
<td>0</td>
<td>( \alpha K )</td>
</tr>
<tr>
<td>PL</td>
<td>( \alpha K )</td>
<td>0  ( \beta(v + \delta D)^- )</td>
<td>( X_2 )</td>
<td>( \alpha K )</td>
<td>0  ( \beta(v + \delta D)^- )</td>
</tr>
</tbody>
</table>

In Table II.2, the \( \alpha K \)'s on date 1 in states GH and BH follow from the same argument in the proof of lemma 2. Notice that if a contract form terminates the project on date 1, the lender can seize \( \alpha K \) only. If a contract form allows the project to continue to date 2, \( X_2 \) or 0 may be realized depending on the project quality by the setup of the model. These two observations attain us the maximum extractable cash flow levels on
Lemma 3 and lemma 4 and the contract forms themselves determine the numbers on date 1 in states GL and BL in Table II.2.

Table II.2 shows that PHSL contract form provides the same or higher maximum amount of cash to the lender in states GH and BH as the other four contract forms do. By assumptions A1 and A2, the PHSL contract form provides the same or higher maximum extractable cash flows as the other four contract forms in states GL and BL. The following lemma summarize our observation above:

**Lemma 5** Given assumptions A1 and A2, the borrowing capacity of PHSL contract form is at least as high as other feasible contract forms.

**Proof.** Given assumption A1, if A2 holds at equality, the borrowing capacity of PHSL contract form is the same as PHPL contract form; otherwise the borrowing capacity of PHSL contract form is strictly higher than that of PHPL contract form. Given assumption A1, the borrowing capacity of PHSL contract form is strictly higher than the borrowing capacity of GLO, PH and PL contract forms. ■

This lemma rules out the need to consider contract forms other than PHSL for higher initial investment requirements, i.e. larger $K$. This is because the lemma 5 says the PHSL contract form can provide the highest possible borrowing capacity among all available contract forms.

To sort out the debt contract form that generates the highest expected gross revenue for the borrower, we first need to obtain the gross revenue in each state under different contract forms. The gross revenue in state GH is $X_1 + X_2 + C + D$, and in state BH, $X_1 + C + D$. For a debt contract that does not allow the continuation of the project in state L, the gross revenue in state L is $\beta D + \alpha K$. For a debt contract that does not allow the continuation of the project in state H, the gross revenue in state H is
When the continuation in state L is allowed, lemma (3) and lemma (4) can help to calculate the ranges of the gross revenue in state L under different contract forms. The following lemma states the results formally with $GR(\cdot)$ denoting the gross revenue for different states —— GL or BL:

**Lemma 6** Given assumptions A1, A2 and condition (II.1),

a) for a debt contract obtaining separating outcomes in state L,

$$GR(GL) \in \left[ X_2 + C + \beta D, X_2 + C + \beta D(1 + \delta) - \delta v \right];$$

b) for a debt contract obtaining pooling outcomes in state L,

$$GR(GL) \in (X_2 + C + \beta D(1 + \delta) - \delta v, X_2 + C + D]$$

and

$$GR(BL) \in (\beta v + \beta D(1 + \delta), v + D].$$

**Proof.** The differences in the gross revenue come from different levels of $\gamma^L$ in the liquidation of $D$. By lemma 3, $\gamma^L \in \left[ \frac{v}{D}, 1 \right]$ for a PHSL contract. By lemma 4, $\gamma^L \in \left[ 0, \frac{v}{D} + \delta \right)$. Since the liquidation of $D$ destroys value, higher $\gamma^L$ means lower gross revenue. Then, the upper bounds for the gross revenues are obtained by setting $\gamma^L$ to its lowest values in different states and the lower bounds or infimum for the gross revenues are obtained by setting $\gamma^L$ to its highest values or supremum in different states.

Based on lemma 6, it is easy to compile Table II.3 for the comparison of expected gross revenue among different contract forms. Note that $\Psi = X_2 + C$ in the table.

By lemma 1, the borrower will choose the debt contract form that offers the highest expected gross revenue accruing to him. The next lemma provides a very strong sufficient condition to guarantee the expected gross revenue from PHSL contract form
### Table II.3 Gross Revenue to the Borrower

<table>
<thead>
<tr>
<th>Acronym/Date</th>
<th>States [prob.]</th>
<th>GH [ρh]</th>
<th>GL [ρl]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>PHSL</td>
<td>$X_1$</td>
<td>$\Psi + D$</td>
<td>$\Psi + \beta D$, $\Psi + \beta D(1 + \delta) - \delta v$</td>
</tr>
<tr>
<td>PHPL</td>
<td>$X_1$</td>
<td>$\Psi + D$</td>
<td>$(\Psi + \beta D(1 + \delta) - \delta v, \Psi + D)$</td>
</tr>
<tr>
<td>GLO</td>
<td>$X_1 + \beta D + \alpha K$</td>
<td>0</td>
<td>$\Psi + \beta D$, $\Psi + \beta D(1 + \delta) - \delta v$</td>
</tr>
<tr>
<td>PH</td>
<td>$X_1$</td>
<td>$\Psi + D$</td>
<td>$\beta D + \alpha K$</td>
</tr>
<tr>
<td>PL</td>
<td>$X_1 + \beta D + \alpha K$</td>
<td>0</td>
<td>$(\Psi + \beta D(1 + \delta) - \delta v, \Psi + D)$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>States [prob.]</th>
<th>BH [bh]</th>
<th>BL [bl]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acronym/Date</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>PHSL</td>
<td>$X_1$</td>
<td>$C + D$</td>
</tr>
<tr>
<td>PHPL</td>
<td>$X_1$</td>
<td>$C + D$</td>
</tr>
<tr>
<td>GLO</td>
<td>$X_1 + \beta D + \alpha K$</td>
<td>0</td>
</tr>
<tr>
<td>PH</td>
<td>$X_1$</td>
<td>$C + D$</td>
</tr>
<tr>
<td>PL</td>
<td>$X_1 + \beta D + \alpha K$</td>
<td>0</td>
</tr>
</tbody>
</table>

**Note:** $\Psi = X_2 + C$

is the highest.

**Lemma 7** Assuming $A1$ and $A2$, condition $\alpha K > \nu + (\delta/b)D$ can ensure that the expected gross revenue of PHSL contract form is the highest among those of other available contract forms.

**Proof.** For the expected gross revenue to PHSL contract form to be higher than that of the PHPL form, it suffices to have the minimum expected gross revenue from PHSL be higher than the maximum expected gross revenue from PHPL. That is,

$$
pl(X_2 + C + \beta D) + bl(\beta D + \alpha K) > pl(X_2 + C + D) + bl(\nu + D)
$$

$$
\Rightarrow \ l\beta D + bl\alpha K > bl\nu + ld.
$$

This yields the condition $\alpha K > \nu + (\delta/b)D$.

The gross revenue to GLO is dominated by PHSL: PHSL provides higher gross revenue than GLO in states GH and BH. Because PHSL provides more contractible cash
flow in state GH, GLO has to incur more costly liquidation in state GL to support the same size initial investment $K$. This means lower gross revenue to GLO than to PHSL in state GL. In the one state left to compare, BL, the two contract forms provide same level of gross revenue. By a similar argument, one can show that PL is dominated by PHPL. Thus to show PL is dominated by PHSL, it suffices to have shown that PHPL is dominated by PHSL in gross revenue.

The gross revenue to the PHSL contract form is the same as to the PH contract form in states GH, BH and BL, and is strictly higher in state GL by assumption A1. Thus the expected gross revenue of the PHSL contract form is strictly higher than the PH contract form. This is a consequence of the assumption $X_2 > \alpha K$.

Lemma 5 and lemma 7 support the following result:

**Proposition 8** Given $\min\{X_1, C, X_2\} > \alpha K \geq \beta D \geq v$, $\alpha K > v + (\delta/b)D$ and $K \leq \frac{\rho(X_2+\beta D)}{1-\alpha(1-\rho l)}$, the project will be funded on date 0 and the debt contract form attains pooling outcomes in state H and separating outcomes in state L.

**Proof.** By lemma 7, the PHSL debt contract form satisfies the two properties in lemma 1 under the conditions listed above. Lemma 5 and the breakeven condition for the lender, (II.6) require that $K \leq \frac{\rho(X_2+\beta D)}{1-\alpha(1-\rho l)}$. Proposition 8 shows that there exists a subset of the parameter space that supports the equilibria where the PHSL contract form is optimal. We can further generalize proposition 8 by relaxing the assumption that the cash flow in state L is zero. Let $X_1^H$ denote the first period cash flow in state H and $X_1^L$ denote the first period cash flow in state L. We then need to rewrite assumption A1 into

$$\min\{X_1^H, C, X_2\} > \alpha K$$  \hspace{1cm} (II.20)
and condition (II.16) into

\[ \alpha K \geq \beta(v + \delta D) + X^L_1 \]  

(II.21)

Condition (II.21) simply means the separating outcomes are still attainable given higher than zero low cash flow level. Also notice that condition (II.17) implies

\[ R^L_1 \leq \beta(C + \delta D) + X^L_1. \]

But again, this condition is satisfied by assumption A2. Thus for the separating outcomes to be attainable and feasible, we need

\[ \beta(v + \delta D) + X^L_1 \leq R^L_1 \leq \min\{\beta D + X^L_1, \alpha K\}. \]  

(II.22)

By realizing that as long as we maintain condition (II.21), the results in lemma 5 and lemma 7 won’t change, the following corollary is obvious:

**Corollary 9** Given \( \min\{X^H_1, C, X_2\} > \alpha K \geq \beta D \geq v, \alpha K > v + (\delta/b)D, \alpha K \geq \beta(v + \delta D) + X^L_1 \) and \( K \leq \min\{\rho X^L_1/(1-\alpha(1-\rho)), \rho X_2/(1-\alpha)\} \), the project will be funded on date 0 and the debt contract form attains pooling outcomes in state H and separating outcomes in state L.

Corollary 9 shows that as long as the separating outcomes are attainable, i.e. \( \alpha K \geq \beta(v + \delta D) + X^L_1 \), a higher than zero first period cash flow in state L will not change the optimality of the PHSL debt contract under assumptions A1, A2 and \( \alpha K > v + (\delta/b)D \).
II.D Discussion

II.D.1 Intuitions behind the PHSL Debt Contract

The analysis in the last section provides sufficient conditions to achieve equilibria that can support PHSL contract form as the optimal debt contract. The two key assumptions A1 and A2 deliver the desired results. These two assumptions require that $C > v$, i.e. the control rents for different project qualities need to be different. This is the condition that makes separating possible. Also, there needs to be enough cash generated from costly partial liquidation, $\beta D \geq v$, making it feasible to separate. Proposition 8 also states that the liquidation value of the indivisible asset, $\alpha K$, cannot be too low. A very low liquidation value of the indivisible asset means it is more valuable to preserve the divisible asset, $D$, to maturity, and thus pooling outcomes provide higher gross revenue to the borrower in state L. This shows for the interim default and liquidation to be optimal, the value obtained in this kind of liquidation should be high enough. Another reason that we need the liquidation value of the indivisible asset to be high is that if it is too low, renegotiation-proofness will make the separating outcome unattainable. Lastly, the high $X_1$, $C$ and $X_2$ values in assumption A1 eliminate some uninteresting contract forms from consideration. Although lower $X_1$ and $C$ can make separating in state H possible, it is rather unusual for a firm with good track record and a performing loan to be terminated half way. $X_2 > \alpha K$ means the continuation value of a firm is higher than the liquidation value, which is a reasonable assumption. The control rent in state BH can be made lower than $C$ as long as it is still higher than $\alpha K$, the analysis remains the same ceribus paribus.

By giving the high liquidation value of the indivisible asset to the lender in state
BL, the borrower can obtain a higher claim from $X_2$. Since $K > \alpha K \geq R_L^1, R_H^1$, there has to be some payment out of $X_2$ so as to satisfy the breakeven constraint for the lender. Although in reality the managers just use the costly partial liquidation to show their faith in the future of their business, which is less strict than the separating outcomes in the PHSL contract, the signaling game exactly captures this intuition. The separating outcomes in state L also completely eliminate the balloon risk in that state. Since if the project quality is type B, there will be no cash flow on date 2, and the liquidation on date 1 means the lender can recover some value on date 1 before nothing is left on date 2. The balloon risk remains in the model, however, because the PHSL contract allows type BH to continue to date 2. This is a realistic scenario as well since ample short term revenue cannot guarantee the future profitability of the firm.

Both $\beta D$ and $\alpha K$ have multiple functions in this model. $\beta D$ can generate cash in state L and links the private costs to the borrower in partial liquidation to the control rent for him on date 2. $\alpha K$ serves as the lower bound in the renegotiation of the first period repayment in state H, the upper bound in the first period repayment in state L and provides value to the lender when the borrower defaults on date 1. Additional variables can be introduced to the model that can take each role of $\beta D$ and $\alpha K$ individually. Doing this can help to relax some constraints imposed on the relationships between certain variables in the analysis, however the model would become more cumbersome.

These intuitions behind the optimality of the PHSL contract show that the assumptions made to attain the optimality of this contract are reasonable ones. Balloon risk is reduced for firms passing the financial strength test at times of poor short term performance. But, inefficiency is also generated when the borrower engages in costly signaling.
II.D.2 Implications

The PHSL debt contract does not specify the repayment amount in state H and the amount repaid from the second period cash flow of the type G project. Rather a total expected value is required out of these cash flows after deducting the expected cash flows on date 1 in states GL and BL from the initial investment requirement $K$. The indeterminacy of the repayment amount on date 1 is easy to resolve. It is reasonable to assume the date 1 repayment amount in state H is fixed at $aK$. This is acceptable to both the borrower and the lender since both of them are risk neutral. If the lender has stronger preference for a fast repayment path, this will be the equilibrium outcome too. This fixed repayment for the interim can be interpreted as periodic interest payment or a sinking fund provision. When the interim cash flows to the firm are ample, this obligation is fulfilled. When the interim cash flows fall into state L, i.e., the firm is hit by negative liquidity shocks, the parties can agree on a lower interim repayment. If the borrower still doesn’t want to pay this lowered amount, default occurs and the lender steps in to seize the liquidation value of the firm.

The repayments contingent on cash flow states are very close to the empirical findings in Asquith, Gertner, and Scharfstein (1994). Asquith, Gertner, and Scharfstein (1994) shows that in financial distress, many forms of debt restructurings are possible. For example, one particular form of bank restructuring is that the banks loosen the financial constraints on firms by deferring some principal and interest payments. In view of this model, this decision is equivalent to specifying a lower repayment amount in the low cash flow state and requiring higher amount out of the final profits from type G projects, which is the “separating in state L” equilibrium in the PHSL contract. This analogy follows because of the modeling assumptions that affect the necessary separat-
ing repayment in state GL. In general, the necessary repayment to achieve separation can be even higher than the amount specified in state H. That means the banks can also tighten the financial constraints for the firms in financial distress, which is another case found in Asquith, Gertner, and Scharfstein (1994).

The above interpretation of the model shows that the possible interim default in a debt contract is related to the presence of asymmetric information about project quality. Once the fixed repayment schedule cannot be met, the manager will have to face some hard tests from the creditors. The manager has to jump over certain hurdle in order to survive during hard times. To jump the hurdle, a borrower with good projects may pay some costs for certain cash-saving or cash-generating behaviors. These kind of behaviors can be inefficient. This is an well-known result in signaling games. The costly behaviors for the borrower in general include cutting spending, selling some assets and layoffs. The sole purpose for the costs paid for liquidity is to provide credible information about the future solvency of the firm. In this way, debt is given the role of transmitting income fluctuations into the “distressed selling” link in Fisher’s story.

The models of Bernanke and Gertler (1989) and Kiyotaki and Moore (1997) both emphasize the transmission mechanism of debt through the change of borrowing capacity affected by shocks, or the so called ”balance sheet effect”. The dynamic implications of this model also point to the change in borrowing capacity. When the liquidation value of divisible asset, $\beta D$, diminishes under some aggregate liquidity shocks, and in particular, when it falls below the private control rent of a type B project holder in state L, separating outcomes are no longer attainable. When the separating outcomes are impossible, the large value recovered from liquidation of the indivisible asset on date 1 is gone. This will limit the borrowing capacity of the borrower. Lower borrowing capacity will force the borrower to give up some investments which require large initial outlays.
This intuition is formally shown in an earlier version of this paper. Thus, for certain parameter space values, one may also obtain the standard transmission mechanism of debt in macroeconomic analysis.

The necessary shocks to generate the above effects can be as simple as idiosyncratic liquidity fluctuations in the product markets and “distress selling” markets. Aggregate shocks will certainly work too. A snow storm during the holiday season preventing people from going to the department stores, or temporary consumption preference shifts are enough to push an economy with the PHSL debt contract into inefficient liquidation. Technology shocks will not be necessary in a dynamic general equilibrium model based on the PHSL debt contract.

II.E Concluding Remarks

This paper proposes a new modeling device in the dynamic debt contracting context: a divisible asset that can be liquidated at the private costs to the borrower. The upshot of introducing this device is that the firm with long term solvency issues may be forced to default in the interim. The borrower is willing to do so because the higher value obtained in the liquidation of the firm asset can reduce the repayment amount to the lender from the revenues to a high quality firm. An intuitive interpretation of this result is that by meeting the coupon payments or sinking fund provisions, a firm can show the ability to eventually repay the whole debt obligation. The use of costly signaling lowers the balloon risk for the lender in low cash flow states. Inefficiency is also generated while the costly signaling is in effect. This paper illustrates the possibility of using interim repayment in debt contract as a way of signaling project quality, which has been previously considered only in terms of dividends payments or simply the commitment
to debt level. (See de Matos (2001) for detailed literature review.)

Inefficient partial liquidation helps to explain the inefficiency that occurs in recessions during business cycles. Layoff may serve as a good example of the propagation function of debt in this model for macroeconomic analyses. The interaction between credit market frictions and labor market frictions is brought up in this scenario. A theoretical example of the labor contracting inefficiencies has been provided by Ramey and Watson (1997). It will be an interesting exercise to combine their model with the model in this paper so as to bring the two kind of frictions together. Some empirical evidence for this point has already been provided. In Sharpe (1994), the author considers the possibility that high opportunity costs of working capital may force firms to do less labor hoarding during cyclical downturns. In his empirical study, Sharpe found evidence supporting the view that small firms, and highly leveraged firms who have tighter credit constraints, are less prone to hoard labor over the business cycles. Hanka (1998) found that firms with higher debt reduced their employment more often in the 1970s and 80s. As shown in the discussion, better liquidity flexibility in my model can give firms extra means to meet financial obligations, which allows more firms to get funded in the first place. This intuition may explain why the findings in Sharpe (1994) and Hanka (1998) are consistent with the fact that debt financing has become more and more widespread in the United States in the second half of the last century. This paper assumes inefficiency in layoffs or partial liquidation, but the welfare implication is different from the interpretation by Sharpe and Hanka of their findings. Both of these two authors reckon debt related efficiency improvement as the reason for the layoffs by financially constrained leveraged firms. The welfare implication of my model make it more interesting to macroeconomists and thus an important extension of the current model would be to setup a dynamic general equilibrium model to check the macroeconomic effect of “cash
“saving layoff” where liquidity shocks may be propagated by the loss in human capital as firms use layoffs to save cash.

Not all kinds of loss in the partial liquidation for a firm are equivalent to efficiency loss for the whole economy. For example, discount sales of accounts receivable mostly involves wealth transfer from the firm to the receivable buyer, thus this cost for the borrower may not be a cost for the economy as a whole. Other means of gaining liquidity under financial pressure may be cutting back on their perk consumption by the managers. Decrease in perk consumption may be directly reflected in lower selling, general and administrative expenses or indirectly in lower asset balances, for some firm assets are just acquired as perks rather than productive inputs. Whether managers engage in these kinds of activities under financial pressure needs support from empirical evidence. This model also implies that achieving liquidity flexibility even at some cost may enhance the firm’s borrowing capacity. This is another interesting implication of the model yet to be tested empirically.
Chapter III

Financial Pressure and the Dynamics of Agency Costs

III.A Introduction

A large theoretical literature has considered managerial agency costs, such as insufficient effort and excessive perquisite consumption, that arise from asymmetric information and misalignment of incentives. Many models have been developed to analyze the role of debt in restraining agency costs.¹ This restraining role is essentially dynamic, since managers are persuaded to reduce current agency costs by the prospect of greater future rewards. Debt serves to alter this intertemporal trade-off.

Not until recently has direct empirical evidence for the restraining role of debt been provided by researchers. This evidence is mixed at best, especially for results from

¹For examples, see Grossman and Hart (1982), Jensen (1986), Bolton and Scharfstein (1990), Hart and Moore (1995), Hart and Moore (1998), Zwiebel (1996) and the model in Chapter II. On the other hand, Jensen and Meckling (1976), who propose the idea of perquisite consumption as one form of agency costs, emphasized the agency cost of debt in form of the asset substitution, monitoring costs and bankruptcy costs.
using leverage (i.e., the debt to asset ratio) as the measure of debt. These studies do not adequately account for dynamic aspects, however. First, financial factors threaten a manager when the firm’s liquidity is low, but this is not necessarily the same thing as high leverage. More direct measures of liquidity may be better proxies for financial pressure on management. Second, the restraining effect of financial pressure operates only when a manager anticipates future benefits from his or her job. Moreover, in periods following high financial pressure, management may reduce agency costs in order to restore the financial strength of the firm. Thus, temporarily high financial pressure may have persistent effects on firm performance. Existing empirical studies, which rely on cross sections, cannot adequately capture these dynamic influences.

This paper uses dynamic estimation techniques to assess the effects of financial pressure on two efficiency measures linked to agency costs. The first efficiency measure is the SG&A rate, which is defined as selling and general administrative (SG&A) expenses over the beginning balance of net fixed assets. The second efficiency measure is asset turnover, which is defined as sales over the beginning balance of total assets.\(^2\) High SG&A rates and low asset turnover can result from inefficiency in operating and investment decisions due to low managerial effort. Moreover, managers can readily hide their perk consumption in SG&A expenses, while assets purchased for managerial consumption purposes will tend to reduce asset turnover. The financial pressure measure is the difference between interest expenses and EBITDA (earnings before interest, tax, depreciation and amortization), normalized by the beginning balance of pledgeable as-

\(^2\)Singh and Davidson (2003) argue that the SG&A expenses item on the income statement is where the management can hide their perk consumption. Ang, Cole, and Lin (2000) use a similar item — operating expenses from the financial statement of small firm to proxy perk consumption. Both Ang, Cole, and Lin (2000) and Singh and Davidson (2003) use asset turnover as an inverse measure of agency costs due to poor investment decisions and insufficient effort by management which result in lower sales or large amount of assets with low productivity purchased by the management for excessive perk consumption.
Pledgeable assets can serve as collateral in debt finance, and thus they serve as a reasonable indicator for the borrowing capacity of a firm.

In this study, the standard dynamic panel data estimator proposed by Arellano and Bond (1991) is applied to a large panel of manufacturing firms from 1989 to 2003. Firms with persistently high financial distress are excluded from the sample, since these may have little long run attractiveness to managers. The Arellano-Bond estimator controls for the fixed features of a firm, such as size and industry, by regressing on the differenced levels of the variables. Cyclical factors are accounted for by including year dummies in the estimation. A potential problem for including only financial pressure and efficiency measures in the empirical formulation is that simple alternative explanations exist for observed increases in efficiency under high financial pressure. Improvements in fundamentals and growth opportunities may lead to a buildup of debt, inducing a spurious positive relationship between financial pressure and efficiency. To disentangle the reactions of efficiency measures to genuine financial pressure from their reactions to financing behavior, this study applies a standard decomposition technique in multivariable time series analysis, using proxy variables for fundamentals and growth opportunities, in order to identify the pure effect of changes in financial pressure on agency costs.

The pattern of the responses in the SG&A rate found in this paper confirms the key predictions of agency cost theory. A pure surprise increase in financial pressure leads the SG&A rate to decrease significantly for the first two years, and then rebound to levels above its long-run level before the impact finally dies out. In the benchmark

---

3 Similar ideas on how to gauge financial pressure can be found in Asquith, Gertner, and Scharfstein (1994), Nickell and Nicolitsas (1999) and Hovakimian and Titman (2003). Asquith, Gertner, and Scharfstein (1994) use this difference directly without scaling it, Hovakimian and Titman (2003) use the same difference for the identification of firms that are in financial distress and Nickell and Nicolitsas (1999) use the ratio of interest expenses over EBITDA instead of the difference.
case, a doubling of financial pressure relative to its sample mean (amounting to about 55% of its sample standard error) will reduce the SG&A rate by 15 percent per year, relative to its long-run level, for the first two years. The SG&A rate will then exceed its long-run level by an annual average of 4.6 percent for years three through five. The responses of the SG&A rate to the financial pressure shock confirm the theoretical idea that future perk consumption acts as an inducement for management to cut current perks.

The responses of asset turnover are also consistent with agency cost theory. In the case of asset turnover, it is assumed that pure increases in financial pressure affect asset turnover with a one year lag. This assumption is made in order to exclude the effects of changes in fundamentals, as represented by pure changes in asset turnover, from the measured pure changes in financial pressure. A sudden pure increase in financial pressure leads to a positive response of asset turnover for all the years under consideration. In particular, from year two to year five following an unanticipated doubling of financial pressure relative to its mean, asset turnover rises above its long-run level by an annual average of 2.1 percent. It is also worth mentioning that the effect of a pure change in financial pressure on financial pressure itself dies out in two years. Thus the responses in both SG&A rate and asset turnover are very persistent. This is consistent with the idea that managers seek to rebuild financial strength following episodes of high pressure.

Two simple and nonexclusive explanations are helpful in interpreting these results. The first explanation is that managers may exert effort to increase sales while reducing expenditures on selling, general and administrative matters. The possibility of financial pressure to raise effort is thus revealed. The second explanation is that managers may cut unnecessary perk investments and expenses, thus reducing SG&A rate directly and raising asset turnover indirectly by lowering total assets. The author
is aware of no other well formed dynamic models of firm behavior that can provide satisfactory alternative interpretations for these empirical results.

I further test whether the effect of financial pressure on the efficiency ratios is only because that firms may sell some assets in dealing with cash shortfall. The tests are carried out by including sales of long term assets or short term asset balances into the benchmark model. The results show that adding asset sales to the model does not alter the impact of financial pressure on the efficiency ratios. Since interest payments are more stable than cash flow, the effect of financial pressure can be largely attributed to the fluctuation of cash flow. The empirical results prove this point. Thus the effect of other monitoring mechanisms such as board size, outside board members, and pressure from takeover raiders cannot be excluded from consideration, although these mechanisms do not act like debt as a hard constraint on the manager and theories on the dynamic effects of these mechanisms have yet to be provided.

**Related literature.** Ang, Cole, and Lin (2000) first suggest the use of asset turnover and expense ratio to gauge agency costs for outside equity. Their study also investigates the effect of leverage on these efficiency measures. With data from the National Survey of Small Business Finances (NSSBF), Ang, Cole and Lin find that though bank debt can significantly lower the expense ratio, but it is also correlated with lower asset turnover. Leverage, on the other hand, can significantly increase asset turnover, it has an insignificant negative effect the on expense ratio. Singh and Davidson (2003) conduct a similar study using large firm data from the COMPUSTAT database. In their sample, leverage has a significant negative effect on asset turnover, while its negative effect on the expense ratio is not uniformly significant.

In a recent paper based on a large sample of U.K. firms, Florackis and Ozkan (2005) find that bank debt and short-term debt can help to increase asset turnover. Yet
for the most part, the coefficient on the total debt ratio is not statistically significant. The effects of short-term debt are similar to the effects of financial pressure suggested earlier. Both indicate that management will risk their future benefits if the debt obligations cannot be met on time. For the expense ratio, the total leverage of UK firms does have some significant negative effect, but there is no significant effect for bank debt or short-term debt. In another relevant study, Yermack (2005) gathers data on the personal use of company aircraft to measure the degree of agency costs. Again, leverage attains no statistical significance in his regression results. While these results raise questions about theories of debt contracting, they fail to consider the effects of debt in a dynamic context. As argued above, this limits their usefulness as tests of the theories.

The rest of Chapter III proceeds as follows. In Section III.B, I describe the background behind the empirical framework, sample selection, variable construction and econometric specification. Section III.C presents the empirical results of the paper. Section III.D summarizes the findings and discusses their interpretations. Section III.E is an Appendix with detailed estimation results from the Arellano and Bond (1991) GMM estimator.

### III.B Empirical Framework

#### III.B.1 Background

Jensen and Meckling (1976) propose excessive perquisite consumption as one form of agency costs born by outside equity holders. Grossman and Hart (1982) suggest

---

4These researchers mainly focus on the effects of ownership structure, as termed by Jensen and Meckling (1976), on agency costs. Their findings are basically consistent with the theoretical predictions on the relationship between ownership structure and agency costs.

5Supporting evidence concerning the disciplinary role of debt on firms’ labor policies has been found by Sharpe (1994), Hanka (1998) and Nickell and Nicolitsas (1999).
debt as an effective way to restrain management from making inefficient investments that support excessive perquisite consumption. The idea of a close relationship between inefficient investments and perk consumption has been followed by many others in explaining the use of debt in capital structure. These models predict that higher leverage is associated with greater efficiency. On the other hand, Jensen and Meckling (1976) also suggest that debt financing can impose its own agency costs, such as asset substitution. Another well known agency problem associated with debt is the “debt overhang” theory proposed by Myers (1977), in which higher leverage reduce the efficiency of a firm. Thus the theoretical relationship between debt and agency costs can go in either direction. For debt to become an effective way of dissuading management from inefficient investment decisions and/or excessive perk consumption, the key question is really how close a firm is to default or bankruptcy, rather than the debt level itself.

The dynamic relationship between agency costs and financial pressure is considered in both Zwiebel (1996) and Chapter II. Zwiebel (1996) points out that when the future value of job positions for the management is not high enough, e.g., the firm is on the edge of bankruptcy or the CEO is close to retirement, the self-disciplinary function of debt will become weaker. For example, a retiring CEO with low future return from his or her current position may engage in “legacy building.” The flip side of this story is emphasized in Chapter II, where high levels of future perks induce low current agency costs when there is financial pressure. Only when the future control rent is high enough will the managers be willing to give up some current perk consumption in order to avoid default. The broad implication of these models is that in response to financial pressure, the agency costs will be lower than average in the short run and then possibly rebound to above average levels for a time.

---

From the discussions above, it is clear that for leverage to have a disciplinary effect on agency costs, two conditions must be met. First, the disciplinary effect is felt when liquidity is low, not simply when leverage is high. Second, managers respond to financial pressure only when they anticipate receiving future benefits as long as the firm avoids default. In my sample selection, I restrict attention to firms that do not have consistently high financial pressure over the sample period in order to ensure that managers have sufficient prospects for obtaining future benefits from the firm.

III.B.2 Data and Variables

Sample Selection. The financial statement data of sample firms used in this paper are extracted from COMPSTAT. The sample includes publicly traded manufacturing firms (SIC codes between 2000 - 3999) listed on NYSE, AMEX, NASDAQ and regional stock exchanges in the U.S. The data file includes both actively traded companies and companies that no longer exist, so the sample does not have survival bias. The sample period is from 1989 to 2003. Because of the significant change in accounting rules on consolidation requirements in 1989 (FASB94), data before 1989 must be discarded. To avoid accounting irregularities, observations with negative sales are dropped. I also delete firms that tripled in size measured by either assets, sales or employment in any year of the sample. These significant changes may indicate mergers, reorganizations or other major corporate events. Since net fixed assets value is used as an important scaling variable in this paper, firm-years with this variable missing or with a recorded value of zero are deleted. In order to apply the dynamic panel data technique, firms without consecutive observations are dropped. For the purpose of formulating a

---

7See the appendices in Bernanke, Campbell, and Whited (1990) and Gomes (2001).
vector autoregression with at least two lags, I delete firms with fewer than three years of observations. This leaves me an unbalanced panel consisting of 2753 firms with a total of 28599 firm-year observations from 1989 to 2003.

As discussed above, if a firm is in long term financial distress, opportunities for future benefits may be persistently low. Ofek (1993) points out that managers of these firms may choose not to respond to financial pressure. Excluding long-term distressed firms from the sample also reduces the influence of management turnover due to poor performance. Zwiebel (1996) and Chapter II, for examples, both assume it is the same manager who engages in performance-improving behavior under financial pressure. Empirical evidence shows that financially healthy firms do not experience abnormal management turnover. Gilson (1989) finds in a sample of firms with poor performance that 52% of the firms experience a senior-level management change during financial distress, but this rate is only 19% when these firms are not distressed, even though they remain highly unprofitable. In a sample of large firms with poor performance, John, Lang, and Netter (1992) find the turnover rate is actually normal, about 12.5-16.6% per year. Thus, we can reasonably assume that in a sample without financially distressed firm, the senior management team should not experience abnormal turnover and as major agents, they will consider intertemporal optimization of benefits. Furthermore, financial distress may lead to direct increases in SG&A expenses due to anti-takeover measures and reorganization costs, that are not tied to managerial effort or perk consumption. For these reasons, firms with more than half of their observations in financial distress are excluded from the sample. The definition of financial distress comes from Hovakimian and Titman (2003), where a firm is classified as financially distressed in a given year if the firm’s EBITDA is less than 80 percent of its interest expense in one year or if its EBITDA is less than the interest expense for two consecutive
years. After deleting these distressed firms, there are 2418 firms with 25884 firm-years left in my sample.

**Variables.** The variables of interest are asset turnover, growth opportunities, financial pressure and SG&A rate. Asset turnover is defined as the ratio of sales (COMPUSTAT data item 12) in the year to the total book value of firm assets (COMPUSTAT data item 6) at the beginning of the year. As discussed in the Introduction, in order to control for the influences of the financing decisions related to a firm’s growth opportunities, a proxy variable for the growth opportunities should be included in the framework. Two alternative proxies will be considered: price-to-book ratio or PB ratio and the investment rate. The PB ratio is defined as the market capitalization of a firm over the book value of the firm’s common equity (COMPUSTAT data item 60). For the market capitalization value, I extract data from the CRSP database and take the average market capitalization over each year’s last three months as the annual market capitalization for the firm. I then merge this market data with financial statement data from COMPUSTAT using 8-digit CUSIP code. The second measure for growth opportunities is the investment rate, which is defined as the ratio of capital expenditures (COMPUSTAT data item 128) to the beginning-of-the-year net book value of property, plant and equipment (COMPUSTAT data item 8).

The measure for financial pressure in this paper is based on similar measures adopted in Asquith, Gertner, and Scharfstein (1994), Nickell and Nicolitsas (1999), and Hovakimian and Titman (2003). I take the difference between interest expense (COMPUSTAT data item 15) and EBITDA (COMPUSTAT data item 13) and scale it by a proxy variable for the firm’s pledgeable assets at the beginning of the year. Ideally, the PB ratio is a commonly used indicator for growth opportunities. For example, with a simple dividend discount model for equity valuation, the PB ratio can be written as \( \frac{P}{BV} = \frac{E}{r-g} \), which is increasing in the growth rate \( g \) as long as \( r - g > 0 \).
as in Nickell and Nicolitsas (1999), the scaling variable should be the cash flow, i.e., EBITDA itself. The high variability of firms’ performances in my sample, particularly negative cash flow observations, prevents me from doing this. Thus, I use measures of firms’ pledgeable assets to scale financial pressure. There are two proxies for the pledgeable assets. The first is the beginning balance of net fixed assets and the second is a constructed measure that follows Berger, Ofek, and Swary (1996); see also Campello (2005). Berger, Ofek, and Swary (1996) use data on proceeds from discontinued operations in a sample of COMPUSTAT firms over the 1984-1993 period to gauge how much exit value will be discounted from the book value of tangible assets in liquidation. They find that a dollar’s worth of book value produces 72 cents for total receivables, 55 cents for inventories and 54 cents for fixed assets. Thus, the pledgeable assets for a firm may be calculated as:

\[
\text{Pledgeable assets} = 0.715 \times \text{Receivables} + 0.547 \times \text{Inventory} + 0.535 \times \text{Fixed Capital}.
\]

I call the financial pressure measure that is scaled by net fixed assets FP1 and the measure scaled by pledgeable assets FP2. I also calculate the individual EBITDA rate and interest expense rate scaled by the beginning balance of net fixed assets in order to test their effect separately.

I also use the beginning balance of net fixed assets to normalize SG&A expenses (COMPUSTAT data item 189), and I call this ratio the SG&A rate. In doing so, I can avoid a problematic correlation between measured financial pressure and the usual expense ratio.\(^9\) For this reason, I use a more stable number — net fixed assets — to

---

\(^9\)The usual expense ratio is defined as SG&A expenses over sales. Since an increase in financial pressure may be driven by a plunge in sales, a positive relationship between financial pressure and the expense ratio may emerge simply as a consequence of scaling.
normalize SG&A expenses.

One alternative explanation for the effect of financial pressure on asset turnover is the asset sales associated with the liquidity shortfall. To test this hypothesis, I calculated two measures. The first one is the sales of long term assets by adding the Sale of Property, Plant, and Equipment (COMPUSTAT data item 107) and Sale of Investments (COMPUSTAT data item 109). The second one is the balance of short term assets by adding the balance of accounts receivable (COMPUSTAT data item 2) and inventories (COMPUSTAT data item 3). Both sums are normalized by the beginning balance of net fixed assets.

All the variables discussed above are deflated by the CPI and winsorized at 1% and 99%. Summary statistics for the variables in my sample are shown in Table III.1. I also provide some other key financial statement measures in Table III.1, such as leverage, the usual expense ratio and the scaling variables.

Table III.2 shows the distribution of firm life in my sample. About one third of the sample firms have observations for all 15 years during the sample period. Only about one fifth of these firms have fewer than seven observations. Thus, I did not lose a lot of observations as I added lags in the estimated models.

**III.B.3 Econometric Specification**

The Arellano and Bond (1991) generalized method of moments (GMM) estimator is the standard econometric procedure for analyzing large dynamic panel data sets. This technique controls for all effects of fixed firm characteristics and cyclical factors by regressing on differenced variables and including year dummies. Since there are several variables in the analysis, the vector autoregression (VAR) method is used to es-
<table>
<thead>
<tr>
<th>Variable (unit)</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
<th># of Obs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asset Turnover</td>
<td>1.35</td>
<td>0.64</td>
<td>0.16</td>
<td>3.85</td>
<td>23477</td>
</tr>
<tr>
<td>PB ratio</td>
<td>2.47</td>
<td>2.69</td>
<td>0.0095</td>
<td>18.99</td>
<td>22203</td>
</tr>
<tr>
<td>Investment rate</td>
<td>0.27</td>
<td>0.28</td>
<td>0</td>
<td>1.81</td>
<td>23477</td>
</tr>
<tr>
<td>EBITDA rate</td>
<td>0.82</td>
<td>1.31</td>
<td>-7.48</td>
<td>7.74</td>
<td>23381</td>
</tr>
<tr>
<td>Interest Expense rate</td>
<td>0.10</td>
<td>0.17</td>
<td>0</td>
<td>1.41</td>
<td>23477</td>
</tr>
<tr>
<td>FP1</td>
<td>-0.71</td>
<td>1.28</td>
<td>-7.42</td>
<td>8.20</td>
<td>23381</td>
</tr>
<tr>
<td>FP2</td>
<td>-0.35</td>
<td>0.37</td>
<td>-1.69</td>
<td>3.42</td>
<td>23323</td>
</tr>
<tr>
<td>SG&amp;A rate</td>
<td>2.19</td>
<td>3.55</td>
<td>0</td>
<td>26.06</td>
<td>23477</td>
</tr>
<tr>
<td>ARINV rate</td>
<td>2.74</td>
<td>3.69</td>
<td>0.12</td>
<td>24.7</td>
<td>23391</td>
</tr>
<tr>
<td>Leverage(%)</td>
<td>21.12</td>
<td>23.31</td>
<td>0.00</td>
<td>129.34</td>
<td>21739</td>
</tr>
<tr>
<td>SG&amp;A/Sales</td>
<td>0.24</td>
<td>0.16</td>
<td>0</td>
<td>1.53</td>
<td>25884</td>
</tr>
<tr>
<td>Net Fixed Assets (millions)</td>
<td>674.90</td>
<td>3030.27</td>
<td>0.0032</td>
<td>70737.28</td>
<td>25884</td>
</tr>
<tr>
<td>Total Assets (millions)</td>
<td>2024.88</td>
<td>9212.89</td>
<td>0.17</td>
<td>302254.7</td>
<td>25883</td>
</tr>
</tbody>
</table>

**Definitions of the Variables**

- Asset Turnover = Sales$_t$/Total Assets$_{t-1}$
- PB ratio = Market Capitalization$_t$/Book Value of Common Equity$_t$
- Investment rate = Capital Expenditure$_t$/Net Fixed Assets$_{t-1}$
- EBITDA rate = EBITDA$_t$/Net Fixed Assets$_{t-1}$
- Interest Expense rate = Interest Expenses$_t$/Net Fixed Assets$_{t-1}$
- FP1 = Financial Pressure$_t$/Net Fixed Assets$_{t-1}$
- FP2 = Financial Pressure$_t$/Pledgeable Assets$_{t-1}$
- SG&A rate = SG&A$_t$/Net Fixed Assets$_{t-1}$
- ARINV rate = (Accounts Receivable$_t$ + Inventories$_t$)/Net Fixed Assets$_{t-1}$
- Leverage = Long-term Debt$_t$/Total Assets$_t$

In order to assess the effect of financial pressure on the efficiency variables, the changes in financial pressure must be decomposed into pure changes, reflecting genuine financial pressure, and changes induced by the responses of financial policy to underlying fundamentals and growth opportunities. For example, a new growth opportunity.

\(^{10}\)Within the corporate finance literature, this kind of panel data VAR framework has been adopted by Whited (1992), Gilchrist and Himmelberg (1998), and Gilchrist, Himmelberg, and Huberman (2005), among others.
Table III.2 Distribution of Firm Life

<table>
<thead>
<tr>
<th>Observations of a Firm</th>
<th>Frequency</th>
<th>Percent</th>
<th>Cumulative Freq. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>151</td>
<td>6.24</td>
<td>6.24</td>
</tr>
<tr>
<td>5</td>
<td>186</td>
<td>7.69</td>
<td>13.94</td>
</tr>
<tr>
<td>6</td>
<td>177</td>
<td>7.32</td>
<td>21.26</td>
</tr>
<tr>
<td>7</td>
<td>125</td>
<td>5.17</td>
<td>26.43</td>
</tr>
<tr>
<td>8</td>
<td>171</td>
<td>7.07</td>
<td>33.5</td>
</tr>
<tr>
<td>9</td>
<td>171</td>
<td>7.07</td>
<td>40.57</td>
</tr>
<tr>
<td>10</td>
<td>169</td>
<td>6.99</td>
<td>47.56</td>
</tr>
<tr>
<td>11</td>
<td>158</td>
<td>6.53</td>
<td>54.09</td>
</tr>
<tr>
<td>12</td>
<td>108</td>
<td>4.47</td>
<td>58.56</td>
</tr>
<tr>
<td>13</td>
<td>73</td>
<td>3.02</td>
<td>61.58</td>
</tr>
<tr>
<td>14</td>
<td>102</td>
<td>4.22</td>
<td>65.8</td>
</tr>
<tr>
<td>15</td>
<td>827</td>
<td>34.2</td>
<td>100</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>2418</strong></td>
<td><strong>100</strong></td>
<td></td>
</tr>
</tbody>
</table>

may lead to issuance of new debt, which may raise the financial pressure variable in the short run even though there has been no actual increase in financial pressure.

To control for these sources of spurious variation in the financial pressure variable, I assume that changes in the variables have the following contemporaneous relationship for the $i^{th}$ firm in a given year $t$:

\[
\begin{bmatrix}
  v_{ast}^{it} \\
v_{go}^{it} \\
v_{fp}^{it} \\
v_{SGA}^{it}
\end{bmatrix}
= \begin{bmatrix}
  a_{21}u_{ast}^{it} + u_{go}^{it} \\
  a_{31}u_{ast}^{it} + a_{32}u_{go}^{it} + u_{fp}^{it} \\
  a_{41}u_{ast}^{it} + a_{42}u_{go}^{it} + a_{43}u_{fp}^{it} + u_{SGA}^{it}
\end{bmatrix}. \tag{III.1}
\]

The $v_{it}^j$'s are the contemporaneous changes in each variable that are uncorrelated with past changes in the same variable. The $u_{it}^j$'s are the pure changes, or impulses, in each variable in the current period $t$. These pure changes are assumed to be mutually uncorrelated, i.e., $E(u_{it}^j u_{it}^k) = 0$ for all $j$ and $k$, $j \neq k$. The superscripts $ast$, $go$, $fp$ and $SGA$ refer to asset turnover, growth opportunities, financial pressure and the SG&A
rate, respectively.

In decomposition (1), the impulse to asset turnover, $u_{it}^{ast}$, reflects changes in the firm’s underlying fundamentals in period $t$ that are driving the entire system. This means that the current fundamental will affect all the other variables in the VAR system in the current period. Impulses to growth opportunities, $u_{it}^{go}$, are identified as the part of the current period changes in growth opportunities that cannot be explained by changes in asset turnover. This is implied by the fact that my proxy variables for growth opportunities are affected by fundamentals.

The key identifying restriction is the one that defines impulses to financial pressure, $u_{it}^{fp}$. These are defined as the part of the current period changes in financial pressure that cannot be explained by changes in fundamentals and growth opportunities. This specification removes the possibility that a change in the financial pressure variable induced by fundamentals or growth opportunities is mistakenly identified as a true change in financial pressure. The restriction is very strong, since the omitted part of the variation in the financial pressure variable (i.e., $a_{31}u_{it}^{ast} + a_{32}u_{it}^{go}$) is likely to reflect at least some amount of true financial pressure. But by proceeding in this way we can be confident that the results obtained using the impulses do not reflect spurious correlation.

The assumptions in decomposition (III.1) give only the contemporaneous relationships between the variables. To show the impact of financial pressure on efficiency over time, the estimated VAR coefficients obtained from the Arellano-Bond estimator are combined with the identified contemporaneous impulses responses from (III.1); see the Appendix for technical details. We obtain the following expressions for the SG&A
rate and asset turnover variables:

\[ y_{SGA}^{it} - y_{SGA}^{i} = a_{41}u_{ast}^{it} + a_{42}u_{g0}^{it} + a_{43}u_{fp}^{it} + u_{SGA}^{it} + \sum_{p=1}^{\infty} \psi_{1,p}u_{ast}^{it-p} + \sum_{p=1}^{\infty} \psi_{2,p}u_{g0}^{it-p} + \sum_{p=1}^{\infty} \psi_{3,p}u_{fp}^{it-p} + \sum_{p=1}^{\infty} \psi_{4,p}u_{SGA}^{it-p}, \]  

(III.2)

\[ y_{ast}^{it} - y_{ast}^{i} = u_{ast}^{it} + \sum_{p=1}^{\infty} \phi_{1,p}u_{ast}^{it-p} + \sum_{p=1}^{\infty} \phi_{2,p}u_{g0}^{it-p} + \sum_{p=1}^{\infty} \phi_{3,p}u_{fp}^{it-p} + \sum_{p=1}^{\infty} \phi_{4,p}u_{SGA}^{it-p}, \]  

(III.3)

where \( y_{SGA}^{it} \) and \( y_{ast}^{it} \) denote the SG&A rate and asset turnover, respectively, for the \( i^{th} \) firm in year \( t \). In equation (III.2), a shock to financial pressure in year 0, associated with an impulse \( u_{fp}^{it} \neq 0 \), has a contemporaneous effect on the SG&A rate in year 0, indicated by the coefficient \( a_{43} \). The year 0 shock also has an effect on the SG&A rate in future years, indicated by the coefficients \( \psi_{3,1}, \psi_{3,2}, \psi_{3,3},... \) for years 1, 2, 3, ... This dynamic pattern of effects is called the impulse response function.

A similar interpretation holds for equation (III.3). In this case, the shock to financial pressure in year 0 cannot affect asset turnover in year 0, based on my identifying assumption that contemporaneous changes in asset turnover reflect underlying fundamentals rather than financial pressure. Asset turnover may be affected in subsequent years, however, and the impulse response function of asset turnover to a financial pressure shock is determined by the coefficients \( \phi_{3,1}, \phi_{3,2}, \phi_{3,3},... \) for years 1, 2, 3, ... .

In decomposition (III.1), the assumption that changes in financial pressure will not reflect impulses to the SG&A rate might be questioned on the grounds that sudden changes in SG&A items might be spuriously identified as changes in financial pressure. For example, an increase in selling costs may induce management to borrow in the short term, raising interest payments and thus increasing measured financial pressure.
While it is reasonable to view these changes as perfectly legitimate financial pressure impulses, it still seems worthwhile to check the robustness of the results to this part of the identification assumptions. Thus, effects of financial pressure impulses are also assessed under the following alternative decomposition:

\[
\begin{bmatrix}
v_{it}^{ast} \\
v_{it}^{go} \\
v_{it}^{SGA} \\
v_{it}^{fp}
\end{bmatrix} =
\begin{bmatrix}
v_{it}^{ast} \\
a_{21}v_{it}^{ast} + u_{it}^{go} \\
a_{41}v_{it}^{ast} + a_{42}u_{it}^{go} + u_{it}^{SGA} \\
a_{31}v_{it}^{ast} + a_{32}u_{it}^{go} + a_{33}u_{it}^{SGA} + u_{it}^{fp}
\end{bmatrix}.
\tag{III.4}
\]

This decomposition determines a different set of financial pressure impulses \( u_{it}^{fp} \), based on removing the effects of impulses to the SG&A rate in addition to asset turnover and growth opportunities. The coefficients in equations (III.2) and (III.3) are altered, leading to different estimates of the dynamic effects of financial pressure impulses. For example, the coefficient \( a_{43} \) in (III.2) is now set to zero.

One basic assumption in the econometric framework above is that all firms in my sample behave in the same manner over time. This is why all the coefficients in equations from (III.1) to (III.4) have no individual or time subscripts. Also, after controlling for the fixed time effect, I assume no cross sectional dependence amongst the firms. Thus, in analyzing the empirical results, we are discussing the behavior of a representative firm from the sample.

### III.C Empirical Results

The econometric procedure used in analyzing all cases in the paper is described as follows. I first estimate the VAR system with the Arellano and Bond (1991) estimator using a lag structure that ensures no autocorrelation in the differenced error terms.
Detailed estimation results are given in the Appendix. With the estimated coefficients of the VAR system, the estimated contemporaneous changes \( v_{it} \)'s for each variable are calculated. Then the variance-covariance matrix of the contemporaneous changes in each variable for a representative firm is obtained through averaging the variance-covariance matrices estimated from the \( v_{it} \)'s. Applying the standard Cholesky decomposition to the representative variance-covariance matrix, we can obtain estimated standard deviations for the impulses for each variable associated with a representative firm.

To analyze the impulses, I first apply the variance decomposition technique; details are given in the Appendix. The variance decomposition indicates the proportion of the total variance of forecasting errors for a variable of interest that can be attributed to the forecasting errors for each variable in the VAR system. The larger that proportion is, the more influence the variable has upon the variable of interest. Following this, I consider the impulse response functions for the SG&A rate and asset turnover with respect to an impulse to financial pressure. The calculation of impulse response functions is based on the autoregressive coefficients obtained in the Arellano and Bond (1991) GMM estimation.\(^{11}\) The impulse response functions give the directions and magnitudes of impacts due to financial pressure impulses on the efficiency measures.

### III.C.1 Benchmark Case

In the benchmark case, the PB ratio and FP1 are chosen to represent growth opportunities and financial pressure, respectively.

The first column in Table III.3 shows the magnitudes of the standard deviations of the impulses for each variable. The variance decomposition is shown in columns two

---

\(^{11}\)For a description of the procedure for deriving impulse response functions from autoregressive coefficients, see standard texts such as Lütkepohl (1991) or Hamilton (1994). Chapter IV also contains a brief description of the procedure.
Table III.3 Analysis of Impulses — Benchmark Case

<table>
<thead>
<tr>
<th>Variables</th>
<th>Std. Dev.</th>
<th>Asset Turnover</th>
<th>PB ratio</th>
<th>FP1</th>
<th>SG&amp;A rate</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asset Turnover</td>
<td>0.27</td>
<td>86.60%</td>
<td>10.74%</td>
<td>2.33%</td>
<td>0.34%</td>
<td>100%</td>
</tr>
<tr>
<td>PB Ratio</td>
<td>1.38</td>
<td>2.74%</td>
<td>96.82%</td>
<td>0.29%</td>
<td>0.15%</td>
<td>100%</td>
</tr>
<tr>
<td>FP1</td>
<td>0.56</td>
<td>17.94%</td>
<td>11.64%</td>
<td>69.18%</td>
<td>1.23%</td>
<td>100%</td>
</tr>
<tr>
<td>SG&amp;A rate</td>
<td>0.99</td>
<td>6.74%</td>
<td>2.10%</td>
<td>12.42%</td>
<td>78.74%</td>
<td>100%</td>
</tr>
</tbody>
</table>

through five of the same table. The fourth column shows the proportion of the total variance of each variable that can be attributed to the impulses to financial pressure. For asset turnover, financial pressure impulses account for about 2.33% its total variation. Note that the major variation in asset turnover is caused by the innovations in the variable itself. Impulses to financial pressure are about half as important as impulses to growth opportunities in explaining asset turnover. For the SG&A rate, financial pressure impulses explain about 12.42% of its total variation.

Figure III.1 The Impulse Response Functions to Shocks in FP1

Figure III.1 shows impulse response functions for an impulse to FP1 over a 10 year horizon for the financial pressure variable itself and the two efficiency measures. Also shown are error bands at the 95 percent confidence level. The magnitude of the
impulse is equal to the absolute value of sample mean, reported as 0.71 in Table III.1, which is 127 percent of the estimated standard deviation of the impulse as shown in Table III.3, but only 55% of its sample standard deviation in Table III.1. From the left to the right in Figure III.1, the impulse response functions are for FP1, the SG&A rate and asset turnover, normalized by the sample averages for the corresponding variables. The responses in each variable over time thus are expressed as a percentage of the sample average of that variable. The zero levels on the vertical axes indicate the long-run or steady state values for each variable. Thus the impulse response functions show how each variable deviates from its respective steady state value due to the impulse to FP1.

Consider the impulse response function of the SG&A rate in the second graph in Figure III.1. The initial response in the SG&A rate to a financial pressure impulse is significantly negative, and this response lasts for another year before the variable rebounds above its steady state value. From year three to year five, this response stays at a level that is significantly higher than the steady state value. As for the point estimate: the lowest response level in the SG&A rate occurs one year after the initial impact from financial pressure, where it lies at 15.2 percent below its steady state value. The 95 percent confidence intervals lies within \([-19\%, -12\%]\) for the first two years. Three years after the shock, the response reaches its highest level at 6.8 percent above its steady state value and its 95 confidence interval is \([4.5\%, 9\%]\). The average negative response in the first two years for the SG&A rate is 15 percent, and the average positive response during year 3 to year 5 is 4.6 percent, relative to the steady state.

Next consider the asset turnover impulse response function, shown in the third graph in Figure III.1. Since I assume there is no contemporaneous effect from financial pressure on asset turnover, the impulse response function of asset turnover starts from
zero. Observe that asset turnover responses to the financial pressure shock are significantly positive during year two through year five. Two years after the shock, the asset turnover response reaches its highest level at around 2.7 percent above its steady state value, and the confidence interval for this peak is [1.4%, 4.1%]. During the second year through the fifth year after the shock, asset turnover is increased by about 2.1 percent per year on average. These results indicate that managers either effectively increase sales with the same amount of capital, or achieve the same sales level with fewer total assets. Higher sales is consistent with the idea that managers raise their effort level under financial pressure. Achieving the same sales level with fewer assets is consistent with the idea that managers release previously invested capital that was mainly for perk consumption.

Finally, note that the impact of a financial pressure shock on the financial pressure variable itself (first graph from the left) falls quickly below zero in only two years, while the responses of asset turnover and the SG&A rate both last beyond five years. This shows the persistence of the effects of financial pressure on these two efficiency measures.

The combination of these results yields strong evidence in favor of agency theories. Two years after the financial pressure shock, the SG&A rate starts to turn above its steady state value. This timing coincides with the return of financial pressure to its steady state value. This pattern suggests that managers are willing to cut expenses under the impact of financial pressure, but after the crisis is over, they will enjoy their perks again, and at an above-normal level that helps offset their losses during the previous difficult periods. This pattern confirms the prediction that managers respond to financial pressure by lowering short run agency costs in exchange for long term benefits. Furthermore, the fact that managers are able to cut expenses without sacrificing improvements in asset turnover in the first two years indicates that these reduced expenses were not
primarily productive, at least in the short run. In other words, this strongly indicates that perquisite consumption is hidden in the SG&A item.

### III.C.2 Robustness

Now I check the robustness of results obtained in the benchmark case by examining alternative cases. In the first alternative case, the investment rate is used as a proxy for growth opportunities. In the second alternative case, FP2 is used as the measure of financial pressure. The last alternative case uses the decomposition (III.4) to identify the impulses to financial pressure.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Std. Dev.</th>
<th>Variance Decomposition over 10-year Horizon</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Asset Turnover</td>
</tr>
<tr>
<td>Asset Turnover</td>
<td>0.27</td>
<td>96.69%</td>
</tr>
<tr>
<td>Investment Rate</td>
<td>0.16</td>
<td>14.52%</td>
</tr>
<tr>
<td>FP1</td>
<td>0.55</td>
<td>22.39%</td>
</tr>
<tr>
<td>SG&amp;A Rate</td>
<td>0.92</td>
<td>7.29%</td>
</tr>
</tbody>
</table>

Table III.4 contains the variance decomposition for the first alternative case. When growth opportunities are accounted for by the investment rate, the proportion of the variation in asset turnover attributed to financial pressure impulses is roughly the same compared to the benchmark value. The variation in the SG&A rate explained by financial pressure impulses is about three-fifth of the benchmark value, since the investment rate impulses do a much better job explaining SG&A rate variation. Financial pressure impulses continue to have important explanatory power, however.

Table III.5 for the second alternative case shows that the impulses to FP2 explain more than twice as much of the total variation of asset turnover as in the preceding
Table III.5 Analysis of Impulses — Alternative Case II

<table>
<thead>
<tr>
<th>Variables</th>
<th>Std. Dev.</th>
<th>Variance Decomposition over 10-year Horizon</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Asset Turnover</td>
</tr>
<tr>
<td>Asset Turnover</td>
<td>0.27</td>
<td>86.62%</td>
</tr>
<tr>
<td>PB Ratio</td>
<td>1.37</td>
<td>2.45%</td>
</tr>
<tr>
<td>FP2</td>
<td>0.16</td>
<td>22.03%</td>
</tr>
<tr>
<td>SG&amp;A rate</td>
<td>1.03</td>
<td>7.83%</td>
</tr>
</tbody>
</table>

cases. The proportion of the total variation in the SG&A rate that can be attributed to the impulses to FP2 is lower than the other cases, since the SG&A rate explains a larger proportion of its own variation. But FP2 still attains a higher proportion than the PB ratio in the total variation of the SG&A rate.

Table III.6 Analysis of Impulses — Alternative Case III

<table>
<thead>
<tr>
<th>Variables</th>
<th>Std. Dev.</th>
<th>Variance Decomposition over 10-year Horizon</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Asset Turnover</td>
</tr>
<tr>
<td>Asset Turnover</td>
<td>0.27</td>
<td>84.84%</td>
</tr>
<tr>
<td>PB Ratio</td>
<td>1.38</td>
<td>2.78%</td>
</tr>
<tr>
<td>SG&amp;A rate</td>
<td>1.04</td>
<td>6.23%</td>
</tr>
<tr>
<td>FP1</td>
<td>0.55</td>
<td>17.08%</td>
</tr>
</tbody>
</table>

Table III.6 corresponds to the alternative case with a different identification scheme. Since in the decomposition (III.4) the financial pressure impulses in the benchmark case will be partially absorbed by SG&A rate impulses, the proportion of the total variation in asset turnover and the SG&A rate that can be attributed to the financial pressure impulses should be expected to decrease. In Table III.6, this proportion for asset turnover is even a little higher than the benchmark case, however the difference isn’t significant. As Table III.6 shows, the explanatory power of financial pressure impulses for the SG&A rate is reduced relative to the benchmark, but the proportion of variation explained remains important.
a. Investment Rate as the Proxy for Growth Opportunities

b. Financial Pressure Constructed with Pledgeable Assets

c. Alternative Identification Assumptions

Figure III.2 The Impulse Response Functions in Alternative Cases
Figure III.2 shows the impulse response functions of interest in all three alternative cases. They are calculated in the same way as for the impulse response functions in Figure III.1. The first column in Figure III.2 gives the impulse response functions for financial pressure measures, the second column for the SG&A rates and the third column for asset turnover. The horizontal panels a, b and c correspond to the three alternative cases. Each panel in Figure III.2 differs only slightly from Figure III.1, which indicates that the major conclusions from the benchmark case remain valid in all three alternative cases. From left to right in Figure III.2, the responses in financial pressure to its own impulses all die out after two years; the responses in the SG&A rate all follow the drop-and-rebound pattern of the benchmark case, though a difference appears in panel (c) due to the alternative identification assumption; and the responses in asset turnover all remain above long-run level from year two to year five.

The differences between the first alternative case and the benchmark case do not appear to be significant. There are some noticeable differences from the benchmark in the remaining two alternative cases, however. For the second alternative case, the impulse under consideration is much smaller than in the other cases, since the sample mean of FP2 is smaller than the sample mean of FP1. When using FP2 as the financial pressure proxy, the initial drop in the SG&A rate is significantly negative only during the first year after the shock. After the second year, the point estimates of the SG&A rate responses are higher than the benchmark case for all periods. Notably the response of SG&A rate reaches its peak at 11.2 percent (95 percent confidence interval [6.3%, 16.1%]) in the fourth year after the shock, which is about 6.5 percent above the response in the fourth year for the benchmark. But the most dramatic difference from the benchmark is for the responses in asset turnover. The impact of the impulse in FP2 can cause asset turnover to rise above its long-run expectation at significant level even after year 6. During year two
to year five after the initial impulse from FP2, the annual average of the point estimates of the excess asset turnover is 5.4 percent above the steady state value, with a maximum in the second year of 5.8 percent (95 percent confidence interval [2.4%, 9.1%]). Thus, the impulses from FP2 seem to be more powerful than those from FP1 in generating responses in asset turnover, as not only is the sample mean of FP2 smaller, but the standard deviation of its impulses is also smaller, as may be seen by comparing Table III.5 with Tables Table III.3, Table III.4 and Table III.6. The different identification assumption in the third case leads the initial response in SG&A rate to equal zero. This is the only significant difference between the third alternative case and the benchmark.

III.C.3 Tests of Alternative Hypotheses

Two Tests on the Effect of Asset Sales

One alternative explanation for the responses in asset turnover following a financial pressure shock is that the firms may sell some assets to generate some liquidity to repay debt. Thus, the increase in asset turnover may simply be due to the reduction in the asset balance from this kind of cash generating sales rather than a decrease of agency costs. If this alternative hypothesis is true, adding asset sales into the VAR system will lead the positive effect in the asset turnover to disappear from the responses to financial pressure shocks and reappear in the responses to asset sales shocks. Also, a straightforward consequence of this alternative hypothesis is that asset sales should go up after the financial pressure shock.\(^{12}\) To test this alternative hypothesis, I apply two alternative proxies in the empirical formulation. The first one is the asset sales measured by the sum of sale of fixed assets and investments. This is a proxy for the sales of long term

\(^{12}\)I thank Professor Roger Gordon from UCSD and Dr. Yuanzhi Luo from OTA Asset Management for the suggesting these tests.
assets. The second proxy is the sum of balances of receivables and inventories, which represents the changes in short term assets. There is one problem with the use of short term asset balance worth noticing: it is only a stock measure, not a flow measure. But since the cash flow from selling these assets belongs to the operating cash flow generated under sales and is not reported separately, no other feasible proxy is available at this point. However, significant and persistent responses in the short term asset balances can still provide us some information about firm behavior although at any given point in time, the balance carried on a firm’s balance sheet may be purely random. Figure III.3 shows the impulse response functions when the asset sales variable is added following financial pressure but before the SG&A rate into the VAR system.

Figure III.3 The Impulse Response Functions with Asset Sales in the VAR
The top panel in Figure III.3 contains the impulse responses of the SG&A rate and asset turnover to a financial pressure shock. Comparing to the benchmark results, the basic shape and significance for the responses in SG&A rate and asset turnover do not change in any important way in this five variable system. The first picture on the left in the bottom panel shows that the response in asset sales to the financial pressure shock is hardly different from zero. This is not a surprising result, since in the sample selection process I have chosen the firms with only temporary financial pressure, and they do not have to sell many long term assets in dealing with a liquidity drain. The picture in the right bottom corner in Figure III.3 shows the response in asset turnover to an asset sales shock. Contrary to the alternative explanation, asset turnover does not respond to a positive shock in asset sales in a significantly positive manner. A plausible explanation is that replacement of the assets sold takes place and this normal operation will cancel the reduction in total asset balance due to asset sales.

Figure III.4 shows the same type of impulse response functions as in Figure III.3 except that the proxy for asset balance change is now the sum of balances in accounts receivable and inventories. The responses in SG&A rate and asset turnover to a financial pressure shock remain the same. Thus the effect of financial pressure shocks is preserved. In contrast to the previous formulation, the balance of short term assets actually drops in response to the shock in financial pressure. The magnitude is around 17% below its long run average in the first two years. This significant and persistent impact on short term asset balances represents evidence supporting the view that firms may reduce some assets under financial pressure. However, the picture in the bottom right of Figure III.3 reveals that asset turnover only responds to a positive shock in short term assets in a significantly negative manner in the first year after the initial shock occurs. This means the decrease in short term assets does have a positive impact on asset
Figure III.4 The Impulse Response Functions with ARINV in the VAR turnover. Our analysis of the bottom panel in Figure III.4 actually provides some support to the alternative hypothesis for asset sales. But the top panel shows that even if we take out the impact of the change in short term asset balances, the financial pressure effect remains significant and persistent.

The variance decompositions in these two alternative formulation of the five-variable system, which are not reported here, provide similar results to the benchmark case. The examinations of the effect of asset sales then reject the hypothesis that the significant and persistent responses in asset turnover to financial pressure shocks are only due to the reduction of asset balance because of asset sales.

The above two tests show that for the firms in my sample, the asset balance
adjustments under financial pressure occur in short term assets but not in long term assets. It is reasonable to see no adjustments made in long term asset given my sample selection process as previously argued. The immediate and significant reduction of short term asset balances under financial pressure may bear some liquidity costs, and thus can hurt the performance-based bonus to the managers. This situation is consistent with the story in Chapter II in essence. But in general, short term assets is not viewed as a source for perks to the managers in the literature. Since the impulse responses are to the orthogonalized shocks, the increase in asset turnover under financial pressure is not due to changes in asset balances in this five variables VAR system. As a result, the positive responses in asset turnover to financial pressure should be attributed to the higher effort level from the managers according to the agency theory.

The Test of Cash Flow vs. Interest Expenses

Compared to cash flow, interest expenses are much more steady. This is so because the sales of the firm fluctuate from year to year but the amount of debt that a firm carries and interest rates are relatively stable, at least for the sample period. One can also make this observation from Table III.1: the standard deviation of EBITDA rate is 1.31 compared to the standard deviation of interest expense rate, 0.17. Thus the variations in the financial pressure measure are mainly driven by the variations in cash flow.\(^{13}\) I examine this assertion formally by putting the EBITDA rate in the third place and the interest expense rate in the fourth place in a five-variable VAR system. Both variables are normalized by the beginning balance of net fixed assets.

The results are displayed in Figure III.5. The top panel shows the responses in the SG&A rate and asset turnover to a negative shock in the EBITDA rate. The two

\(^{13}\)I thank Dr. Gerald Garvey from Barclays Global Investors for suggesting this test.
Figure III.5 The Impulse Response Functions to Cash Flow and Interest Expenses Shocks

pictures are very similar to the responses of these two efficiency measures obtained previously. This proves the major results discussed in this paper can be attributed to the variations in firms’ cash flow. The variance decomposition, which is not presented here, shows similar results. The bottom panel of Figure III.5 shows the responses of these two efficiency measures to a positive shock in the interest expenses rate. The initial impact on the SG&A rate from a positive shock in interest expenses is significantly positive and the impact quickly dies out to zero in the following years. The first year impact on asset turnover from a positive shock in interest expenses is significantly negative and is also very short lived. These responses to interest expenses shocks are seemingly inconsistent
with the major results discussed earlier. But it is not surprising to see this pattern in the responses to interest expenses shocks. Increasing interest expenses indicate higher debt levels, which results in higher asset levels. This will cause the asset turnover to decrease as long as the new assets cannot generate sales immediately. The increase in the SG&A rate also reflects the increase in corporate activities associated with the higher asset levels. So far, this test can neither prove nor disprove the hypothesis that the response pattern in the SG&A rate and asset turnover are due to financial pressure.

Given the results in Figure III.5, the impact of cash flow may be due to other corporate governance mechanism than the restraining function of debt. Other corporate governance mechanisms that can restrain managerial agency costs include the size and composition of the board for a company and the pressure from takeover raiders. But there are several reasons why these mechanisms are not good substitutes for the role of debt. First, some of these mechanism may resort to the use of debt as the ultimate controlling device. Zwiebel’s (1996) management entrenchment model is built upon the idea that the managers are using debt to substitute for pressure from takeovers. Boards of directors may also use debt in order to give strict incentives to enforce efficient behavior by the managers. Without this kind of hard device, the monitoring of the board may not be effective since the managers can convince them that the managers alone are not to blame. Secondly, some anti-takeover measures such as golden parachutes can reduce the effectiveness of takeover pressure. The takeover pressure in Jensen’s (1986) free cash flow story applies to firms with ample cash flow rather than firms with liquidity drain. Lastly, I am not aware of any theoretical model for those mechanisms that implies the dynamic effects presented in this paper.

The central hypothesis I am testing is that the dynamic effects in the agency cost measures I found are due to the functioning of debt. To give the strongest rejection of
this hypothesis, one has to show empirical evidence that firms using only debt and not any other monitoring mechanisms do not response to cash flow shocks in the dynamic patterns shown in this paper. Currently I do not have enough information to select such a subsample. The dynamic responses to cash flow in a sample of debt-free firms will also provide useful information. However, to the best of my knowledge, this kind of sample is rarely available in the form of panel data. This situation will make the empirical techniques inapplicable. I also do not have enough information to select a subsample that can test whether a mechanism other than debt can produce the dynamic patterns observed in this section. Even if such subsamples were available, the sample size may be too small to apply the empirical techniques used in this exercise. Thus this task has to be left to future research.

III.D Summary and Discussions

Studies on the effect of ownership structure on agency costs yield inconsistent and sometimes contradictory results for the disciplinary role of debt. The results presented in this paper show that using measures of financial pressure and dynamic panel data techniques, evidence can be obtained to confirm the idea that debt can restrain the perquisite consumption of managers and give them incentives to operate the firm more efficiently. First, financial pressure is a better measure of default or bankruptcy threats, which are the reasons for the effectiveness debt in theoretical models. Second, incorporating time series analysis provides a better description of the responses of management to different operating environments over time. After adopting these two empirical strategies, I obtain findings that are consistent with debt financing theories, and the results are robust across different alternative definitions of variables and decomposition assump-
The basic findings include: asset turnover as an inverse measure of agency costs will be higher after an unexpected increase in financial pressure; the rate of SG&A expenses will be lower immediately after an unexpected increase in financial pressure, but it will rise again as the performance improves over time. The simultaneous increase in asset turnover and decrease in SG&A expenses under financial pressure strongly indicate the possibility that efficiency could be improved during normal periods, i.e. agency costs exist when financial pressure is absent. If there is no perk consumption hidden in the SG&A expenses, and thus all these expenses are productive inputs, one should be able to observe asset turnover and the SG&A rate move in the same direction when managers are trying to stimulate sales. The fact that under financial pressure, SG&A expenses actually fall while asset turnover is rising indicate this item indeed contains some unproductive perks consumed by the management, or else managers are able to exert effort to raise turnover while reducing expenses. The persistence of higher SG&A rates long after the financial pressure shock takes place is consistent with the idea that management recognizes dynamic trade-offs when making short-term choices. This intertemporal substitution effect is a unique finding of this paper. No evidence is found to support alternative hypotheses other than the dynamic patterns in the agency cost measures are due to the functioning of debt.

The evidence provided in this paper appeals to the positive predictions from agency models. The role of debt in enforcing management discipline is confirmed. However, whether the responses of management to financial pressure help to achieve the highest possible level of economic surplus is subject to interpretation. The way that debt functions is to force the firm to unconditionally generate and pay out cash to its creditors. This feature of a fixed liquidity obligation may be necessary due to the asym-
metric information pertaining to the future profitability of the firm, as in Chapter II. But in that model the means available to a firm to generate liquidity are not economically efficient. In the real world, it is not uncommon to bear some liquidity discount when assets are exchanged for quick cash. This kind of economic loss may be born by the shareholders of the firm or even by management. A reduction in perk consumption can be costly for the management if they have better outside opportunities. The management under financial pressure may focus only on the operating activities that can generate short term liquidity, and forgo those activities that matter for long term profitability. The short term behavior will cost both the management and the shareholders alike. For example, John, Lang, and Netter (1992) find that in response to a performance decline, large firms cut R&D expenditures, which may impact their own future competitive advantages. A similar situation arises when financially distressed firms sell assets to generate short term cash. Shleifer and Vishny (1992) point out that financially distressed firms may have to sell their specialized assets cheaply when the timing of asset sales occurs during an industry-wide downturn. Ramey and Shapiro (2001) and Pulvino (1998) find supporting evidence to this argument in the aerospace industry and the airline industry, respectively. Brown, James, and Mooradian (1994) find that using proceeds from asset sales to repay debt may only benefit the debt holders, while hurting shareholders’ value.

The empirical findings in this paper also highlight the need for more sophisticated dynamic agency models. The drop-and-rebound pattern in perk consumption suggests the possibility that asymmetric information on inefficient operations is revealed for awhile but then managers can hide their perks as performance improves. The process of operating assets to generate income and repay debt involves more periods than most debt financing models assume. Multiple periods leave management with many opportunities to make adjustments according to different levels of random cash flow. If a firm keeps
getting lucky for several periods, management may expand perk consumption or build a larger empire without triggering the attention of creditors or takeover raiders due to asymmetric information. However, if some adverse income shock hits the firm after the “thriving period,” management may cut perk consumption and improve the efficiency of investment in order to keep their positions, provided that there are enough future benefits. But these improvements reveal the fact that inefficiency existed before. Corporate governance may help keep management from reverting back to the inefficient style for some time. Thus the efficient period may last until the next “thriving period” comes or after levitation of financial pressure, when management can utilize their private information to enjoy more perks again. This shows that the leverage effect on agency costs may appear or not appear in different situations over time, and a well-formulated model is necessary to adequately understand the situation.

### III.E Appendix: GMM Estimation Results of the Benchmark Case

The estimation results for the benchmark case using the Arellano and Bond (1991) GMM estimator are given in Table III.7. The estimation results for all other cases are similar in terms of statistical significance and are available upon request. The estimation procedure is thoroughly discussed in the next chapter.

Table III.7 presents the system-of-equations estimation results using the Arellano and Bond (1991) GMM estimator. As shown on the bottom rows of this table, the $p$-values from second order to sixth order autocorrelation tests are all below the 8% (two tailed) significance level. Thus, in all four equations, no autocorrelation statistics
### Table III.7 Arrelano-Bond (1991) System-of-Equations GMM Estimation Results

#### The Benchmark Case

<table>
<thead>
<tr>
<th>Dependent Variables</th>
<th>Asset Turnover&lt;sub&gt;t&lt;/sub&gt;</th>
<th>PB ratio&lt;sub&gt;t&lt;/sub&gt;</th>
<th>FP1&lt;sub&gt;t&lt;/sub&gt;</th>
<th>SG&amp;A rate&lt;sub&gt;t&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff.</td>
<td>z</td>
<td>Coeff.</td>
<td>z</td>
</tr>
<tr>
<td>Asset Turnover&lt;sub&gt;t−1&lt;/sub&gt;</td>
<td>0.483**</td>
<td>23.69</td>
<td>-0.208**</td>
<td>-3.18</td>
</tr>
<tr>
<td>Asset Turnover&lt;sub&gt;t−2&lt;/sub&gt;</td>
<td>0.026*</td>
<td>2.12</td>
<td>-0.242**</td>
<td>-6.75</td>
</tr>
<tr>
<td>Asset Turnover&lt;sub&gt;t−3&lt;/sub&gt;</td>
<td>0.088**</td>
<td>7.64</td>
<td>-0.074*</td>
<td>-2.31</td>
</tr>
<tr>
<td>Asset Turnover&lt;sub&gt;t−4&lt;/sub&gt;</td>
<td>-0.005</td>
<td>-0.39</td>
<td>0.044</td>
<td>1.16</td>
</tr>
<tr>
<td>Asset Turnover&lt;sub&gt;t−5&lt;/sub&gt;</td>
<td>0.049**</td>
<td>4.52</td>
<td>0.023</td>
<td>0.79</td>
</tr>
<tr>
<td>PB ratio&lt;sub&gt;t−1&lt;/sub&gt;</td>
<td>0.071**</td>
<td>14.45</td>
<td>0.355**</td>
<td>20.83</td>
</tr>
<tr>
<td>PB ratio&lt;sub&gt;t−2&lt;/sub&gt;</td>
<td>-0.017**</td>
<td>-6.43</td>
<td>0.015</td>
<td>1.62</td>
</tr>
<tr>
<td>PB ratio&lt;sub&gt;t−3&lt;/sub&gt;</td>
<td>-0.003</td>
<td>-1.34</td>
<td>-0.059**</td>
<td>-6.90</td>
</tr>
<tr>
<td>PB ratio&lt;sub&gt;t−4&lt;/sub&gt;</td>
<td>0.002</td>
<td>1.09</td>
<td>0.047**</td>
<td>5.71</td>
</tr>
<tr>
<td>PB ratio&lt;sub&gt;t−5&lt;/sub&gt;</td>
<td>0.000</td>
<td>0.20</td>
<td>-0.031**</td>
<td>-3.99</td>
</tr>
<tr>
<td>FP1&lt;sub&gt;t−1&lt;/sub&gt;</td>
<td>0.030**</td>
<td>4.46</td>
<td>-0.082**</td>
<td>-4.35</td>
</tr>
<tr>
<td>FP1&lt;sub&gt;t−2&lt;/sub&gt;</td>
<td>0.031**</td>
<td>7.76</td>
<td>0.022</td>
<td>1.95</td>
</tr>
<tr>
<td>FP1&lt;sub&gt;t−3&lt;/sub&gt;</td>
<td>0.009**</td>
<td>2.95</td>
<td>-0.053**</td>
<td>-4.24</td>
</tr>
<tr>
<td>FP1&lt;sub&gt;t−4&lt;/sub&gt;</td>
<td>0.037**</td>
<td>10.68</td>
<td>0.032**</td>
<td>3.16</td>
</tr>
<tr>
<td>FP1&lt;sub&gt;t−5&lt;/sub&gt;</td>
<td>0.000</td>
<td>-0.15</td>
<td>0.051**</td>
<td>4.97</td>
</tr>
<tr>
<td>SG&amp;A rate&lt;sub&gt;t−1&lt;/sub&gt;</td>
<td>-0.008*</td>
<td>-2.30</td>
<td>0.009</td>
<td>0.73</td>
</tr>
<tr>
<td>SG&amp;A rate&lt;sub&gt;t−2&lt;/sub&gt;</td>
<td>0.006*</td>
<td>2.51</td>
<td>0.023**</td>
<td>3.68</td>
</tr>
<tr>
<td>SG&amp;A rate&lt;sub&gt;t−3&lt;/sub&gt;</td>
<td>0.003</td>
<td>1.83</td>
<td>-0.005</td>
<td>-0.80</td>
</tr>
<tr>
<td>SG&amp;A rate&lt;sub&gt;t−4&lt;/sub&gt;</td>
<td>0.011**</td>
<td>6.53</td>
<td>0.021**</td>
<td>3.98</td>
</tr>
<tr>
<td>SG&amp;A rate&lt;sub&gt;t−5&lt;/sub&gt;</td>
<td>0.000</td>
<td>-0.36</td>
<td>0.024**</td>
<td>-5.91</td>
</tr>
</tbody>
</table>

#### Autocorrelation tests

<table>
<thead>
<tr>
<th>Autocorrelation tests</th>
<th>z</th>
<th>p-value</th>
<th>z</th>
<th>p-value</th>
<th>z</th>
<th>p-value</th>
<th>z</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ar(1)</td>
<td>-10.15</td>
<td>100.0%</td>
<td>-4.22</td>
<td>100.0%</td>
<td>-5.62</td>
<td>100.0%</td>
<td>-3.75</td>
<td>100.0%</td>
</tr>
<tr>
<td>ar(2)</td>
<td>-0.17</td>
<td>56.7%</td>
<td>0.22</td>
<td>41.4%</td>
<td>-0.34</td>
<td>63.3%</td>
<td>0.51</td>
<td>30.4%</td>
</tr>
<tr>
<td>ar(3)</td>
<td>1.01</td>
<td>15.6%</td>
<td>-0.93</td>
<td>82.3%</td>
<td>0.54</td>
<td>29.6%</td>
<td>0.65</td>
<td>25.8%</td>
</tr>
<tr>
<td>ar(4)</td>
<td>-1.44</td>
<td>92.6%</td>
<td>0.91</td>
<td>18.1%</td>
<td>0.73</td>
<td>23.1%</td>
<td>-0.91</td>
<td>81.9%</td>
</tr>
<tr>
<td>ar(5)</td>
<td>1.74</td>
<td>4.1%</td>
<td>1.34</td>
<td>9.0%</td>
<td>-0.58</td>
<td>71.9%</td>
<td>0.32</td>
<td>37.3%</td>
</tr>
<tr>
<td>ar(6)</td>
<td>-0.52</td>
<td>69.8%</td>
<td>-0.86</td>
<td>80.6%</td>
<td>-0.37</td>
<td>64.6%</td>
<td>-1.08</td>
<td>86.1%</td>
</tr>
</tbody>
</table>

#### Notes

(i) Total number of observations: 4992; number of firms: 976; Sample period: 1990-2003.

(ii) Coefficient estimates are all two step results from system-of-equations estimation.

(iii) FP1=Financial Pressure/Net Fixed Assets

(iv) Maximum number of lags used as instruments: 9

(v) ** and * indicate statistical significance at the 1% and 5% (two tailed) test levels, respectively.
beyond the first order for the differenced error terms can lead one to reject the null hypothesis that there is no autocorrelation among them.\textsuperscript{14} From the $z$-values are calculated with two-step standard errors, we can see that in this benchmark case, the first four lags of financial pressure have a significant positive coefficients on the asset turnover. For the SG&A rate equation, the first lag of financial pressure has a significant negative coefficient, but the next three lags have significant positive coefficients. These coefficient estimates yield the pattern seen in Figure III.1. For the variance decomposition in the VARs, see Lütkepohl (1991) or Hamilton (1994).

\textsuperscript{14}As can be seen from the formulation, the first order autocorrelation exists by construction and for the past levels of the variables to be valid instruments, we only need no autocorrelations beyond first order, see more details in Chapter IV.
Chapter IV

Asymptotic Distributions of Impulse Response Functions in Short Panel Vector Autoregressions

IVA  Introduction

In this chapter\(^1\), we consider the panel vector autoregressions (VARs) where the cross sectional dimension \((N)\) is large and the time series dimension \((T)\) is short (typically less than 10). Panel VARs with a short \(T\) have been investigated, for example, by Holtz-Eakin, Newey, and Rosen (1988) and Binder, Hsiao, and Pesaran (2005). While these papers focus on the estimation of the slope coefficients, our focus here is on the estimation of the impulse response functions (IRFs) and their confidence bands. Following the traditional panel data literature, we assume that the slope coefficients are the same across different cross sectional units and there is no cross sectional dependence af-

\(^1\)This chapter, in its entirety, is coauthored with professor Yixiao Sun.
ter controlling for the fixed time effects. These two assumptions allow us to make good long-horizon forecasts, especially when the forecasting horizon is comparable to the time series length. This argument is consistent with Binder, Hsiao, and Pesaran (2005) in which short panel VARs are used to infer the long run properties of the underlying time series.

For time series data, VAR models are typically estimated using the equation-by-equation OLS as it is asymptotically equivalent to the full system-of-equations estimator. For panel data VARs, the OLS estimator is inconsistent for a fixed $T$ as $N \to \infty$. In this case, the VAR models are typically estimated using the Anderson and Hsiao (1982) (hereafter AH) estimator or the Arellano and Bond (1991) (hereafter AB) estimator. We can apply these estimators to each equation in the VAR system or the full system of equations. Holtz-Eakin, Newey, and Rosen (1988) pointed out that it may be possible to improve the efficiency by estimating the system of equations jointly. We show that, under the model specification given below, the equation-by-equation AB or AH estimator is asymptotically equivalent to the corresponding system-of-equations estimator.

Impulse response analysis in the time series setting has been provided by Baillie (1987), Lütkepohl (1989) and Lütkepohl (1990), among others. However, there are two important differences between the time series case and the short panel case.

First, for a time series VAR, the OLS estimator of the slope coefficients is asymptotically independent of the covariance estimator while for short panel VARs, the AH or AB estimator of the slope coefficients depends on the covariance estimator even in the limit as $N \to \infty$ for a fixed $T$. Due to the presence of the fixed individual effects, the regressors in the short panel data VAR model are correlated with the regression error. This nonzero correlation leads to the asymptotic dependence between the slope estimator and the covariance estimator.
Second, for a time series VAR, the covariance estimator based on the estimated OLS residual is asymptotically equivalent to that based on the true error term. For the short panel VARs, the covariance estimator has different asymptotic distributions, depending on whether the error term is known or is based on consistent estimates of the slope coefficients. In other words, the estimation uncertainty of the slope coefficients affects the asymptotic distribution of the covariance estimator.

These two differences imply that the usual asymptotic result for orthogonalized impulse responses are not applicable to short panel VARs. The main contribution of the paper is to derive the asymptotic distribution of the orthogonalized IRFs for short panel VARs. Based on our asymptotic result, confidence bands for the IRFs can be easily constructed. Although impulse responses analysis using short panels has been employed in the empirical applications, to the best of our knowledge, no study has reported confidence bands for orthogonalized IRFs that account for the estimation uncertainty of the covariance matrix. As a result, the reported confidence bands are more narrow than they should be. This may lead to the finding of statistical significance that does not actually exist.

Throughout the paper, vec denotes the column stacking operator and vech is the corresponding operator that stacks only the elements on and below the main diagonal. As usual, the Kronecker product is denoted by \( \otimes \), the commutation matrix \( K_{mn} \) is defined such that, for any \((m \times n)\) matrix \( G \), \( K_{mn} \text{vec}(G) = \text{vec}(G') \), and the \( m^2 \times (m(m + 1))/2 \) duplication matrix \( D_m \) is defined such that \( D_{m} \text{vech}(F) = \text{vec}(F) \) for a symmetric \((m \times m)\) matrix \( F \). Furthermore, \( D_m^+ = (D_m'D_m)^{-1}D_m' \) and \( L_m \) is the \((m(m + 1))/2 \times m^2 \) elimination matrix defined such that, for any \((m \times m)\) matrix \( F \), \( \text{vech} (F) = L_m \text{vec}(F) \).

Section IV.B describes the vector autoregressions model for panel data and presents
the GMM estimators of the slope coefficients and analysis-of-variance-type estimator of the error covariance matrix. This section also establishes the joint asymptotic distribution of the coefficients estimator and covariance estimator. Using the asymptotic results in section IV.B, section IV.C derives the asymptotic distributions of the orthogonalized and non-orthogonalized impulse response functions. Sections IV.D provides the simulation evidence showing that the confidence bands derived in section IV.C provide more accurate coverage than the bands constructed without considering the asymptotic dependence of the coefficient estimator and the covariance estimator and the extra estimation uncertainty in the covariance estimator. Section IV.E concludes the paper. The Appendix, section IV.F, provides the proof for the main theorem in section IV.B.

IV.B The Model and Estimation

To construct the impulse responses, we consider an $m$-dimensional $VAR(p)$ process:

\begin{align*}
    w_{i,t} &= c + A_1 w_{i,t-1} + ... + A_p w_{i,t-p} + e_{i,t} \quad (IV.1) \\
    e_{i,t} &= \mu_i + \lambda_t + \epsilon_{i,t} \quad (IV.2)
\end{align*}

for $t = T_{i0}, T_{i0} + 1, ..., T_i$ and $i = 1, 2, ..., N$ where $w_{i,t} = (w_{1,it}, ..., w_{m,it})'$, $A_j$ are $(m \times m)$ coefficient matrices, $\mu_i$, $\lambda_t$, and $\epsilon_{i,t}$ have zero means and are independent among themselves and with each other. We allow for unbalanced panel data sets. For individual $i$, the time series starts at period $T_{i0}$ and ends at period $T_i$. We assume that

\begin{equation}
    E(\epsilon_{i,t} | w_{i,t-1}, ..., w_{i,T_{i0}}) = 0 \text{ for } T_{i0} \leq t \leq T_i \quad (IV.3)
\end{equation}
and for $T_{i0} \leq t \leq s$,

$$E \left( \varepsilon_{i,t} \varepsilon'_{j,s} | w_{i,t-1}, \ldots, w_{i,T_{i0}} \right) = \begin{cases} 
\Sigma, & i = j \text{ and } t = s \\
0, & \text{otherwise}
\end{cases} \quad \text{(IV.4)}$$

where $\Sigma$ is a positive definite matrix.

The model is the same as that considered by Binder, Hsiao, and Pesaran (2005). Here we do not specify the initial conditions for the VAR model as the asymptotic properties of the GMM estimator used in this paper do not depend on how the process is initialized. We focus on the GMM type estimator because it is widely used in empirical applications, see, for example, Love and Zicchino (2002) and Gilchrist, Himmelberg, and Huberman (2005).

Let

$$y_{i,t} = w_{i,t} - w_{.,t}, \quad c_i = \mu_i - \frac{1}{N} \sum_{j=1}^{N} \mu_j, \quad u_{i,t} = \varepsilon_{i,t} - \varepsilon_{.,t} \quad \text{(IV.5)}$$

where $w_{.,t}$ and $\varepsilon_{.,t}$ are the cross sectional averages of $w_{i,t}$ and $\varepsilon_{i,t}$, respectively. Due to the possible unbalancedness, the cross sectional averages may be taken over different numbers of observations for different periods. Writing the VAR system in terms of $y_{i,t}$, we have

$$y_{i,t} = c + A_1 y_{i,t-1} + \ldots + A_p y_{i,t-p} + c_i + u_{i,t}, t = T_{i0} + p, \ldots, T_i. \quad \text{(IV.6)}$$

It is now well known that, due to the correlation between the fixed effect $c_i$ and the regressors, the OLS estimator of $A_i$ based on equation (IV.6) is inconsistent when $T$ is small. To remove the fixed individual effect, we take the first difference of equation (IV.6), leading to

$$\Delta y_{i,t} = A_1 \Delta y_{i,t-1} + \ldots + A_p \Delta y_{i,t-p} + \Delta u_{i,t}, \quad t = T_{i0} + p + 1, \ldots, T_i \quad \text{(IV.7)}$$
The OLS estimator based on the above equation is still inconsistent because $\Delta u_{i,t}$ is correlated with $\Delta y_{i,t-1}$. The standard GMM estimators of AH and AB employ instruments that are orthogonal to $\Delta u_{i,t}$. For the AH estimator, the underlying moment conditions are

$$E \left( \sum_{t=T_{i0}+p+1}^{T_i} \Delta u_{i,t} y_{i,t-1-\ell}' \right) = 0 \text{ for } \ell = 1, 2, \ldots, p; \quad \text{(IV.8)}$$

while for the AB estimator, the moment conditions are

$$E \left( \Delta u_{i,t} y_{i,t-1-\ell}' \right) = 0 \text{ for } \ell = 1, 2, \ldots, t - 1; \ t = T_{i0} + p + 1, \ldots, T. \quad \text{(IV.9)}$$

To write the equations in the vector form, we let

$$\frac{A'}{m \times mp} = (A_1, A_2, \ldots, A_p),$$

and

$$\frac{\Delta y_{i}}{(T_i-T_{i0}-p) \times m} = \begin{pmatrix} \Delta y_{i,T_{i0}+p+1} \\ \Delta y_{i,T_{i0}+p+2} \\ \vdots \\ \Delta y_{i,T_i} \end{pmatrix}, \quad \frac{\Delta u_{i}}{(T_i-T_{i0}-p) \times m} = \begin{pmatrix} \Delta u_{i,T_{i0}+p+1} \\ \Delta u_{i,T_{i0}+p+2} \\ \vdots \\ \Delta u_{i,T_i} \end{pmatrix}, \quad \text{(IV.10)}$$

$$\frac{\Delta X_{i,t}}{mp \times 1} = \begin{pmatrix} \Delta y_{i,t-1} \\ \Delta y_{i,t-2} \\ \vdots \\ \Delta y_{i,t-p} \end{pmatrix}, \quad \frac{\Delta X_{i}}{(T_i-T_{i0}-p) \times mp} = \begin{pmatrix} \Delta X_{i,T_{i0}+p+1} \\ \Delta X_{i,T_{i0}+p+2} \\ \vdots \\ \Delta X_{i,T_i} \end{pmatrix}. \quad \text{(IV.11)}$$

then

$$\text{vec}(\Delta y_{i}) = \text{vec}( (\Delta X_{i}) A ) + \text{vec}(\Delta u_{i}). \quad \text{(IV.12)}$$

In other words,

$$\Delta \tilde{y}_{i} = \Delta \tilde{X}_{i} \alpha + \nu_{i}, \quad \text{(IV.13)}$$
where

\[
\Delta \tilde{y}_i = vec(\Delta y_i), \quad \Delta \tilde{X}_i = (I_m \otimes \Delta X_i) , \quad \alpha = vec(A), \quad v_i = vec(\Delta u_i).
\]

(IV.14)

To construct the instrument matrix, we normalize \( \min(T_i_0) \) to be zero and denote \( T = \max(T_i) \). For the AH estimator, we let \( \tilde{Z}_i = I_m \otimes Z_i \) where

\[
Z_i = \begin{pmatrix}
  y'_{i,0}, \ldots, y'_{i,T_i_0+p-1} \\
y'_{i,1}, \ldots, y'_{i,T_i_0+p} \\
\vdots \\
y'_{i,T_i-p-1}, \ldots, y'_{i,T_i-2}
\end{pmatrix}
\]

is \( m(T_i - T_i_0 - p) \times pm \) matrix with all missing values replaced by zeros. For the AB estimator, we first let

\[
Z^0_i = \begin{pmatrix}
  (y'_{i,0}, \ldots, y'_{i,p-1}) & 0 & \ldots & 0 \\
  0 & (y'_{i,0}, \ldots, y'_{i,p}) & 0 & \ldots \\
  \vdots & \vdots & \vdots & \vdots \\
  0 & 0 & \ldots & (y'_{i,0}, \ldots, y'_{i,T_i-2})
\end{pmatrix}
\]

(IV.16)

which is an \( m(T - p) \times m(T + p - 1) (T - p) / 2 \) matrix. Again the missing values of \( y_{i,t} \) in \( Z^0_i \) are replaced by zeros. Next, we keep the rows of \( Z^0_i \) that correspond to periods \( T_i_0 + p + 1, \ldots, T_i \) and remove the rest of the rows. We denote the resulting matrix as \( Z_i \), an \( m(T_i - T_i_0 - p) \times m(T + p - 1) (T - p) / 2 \) matrix and define \( \tilde{Z}_i = I_m \otimes Z_i \).

With these definitions of \( \tilde{Z}_i \), the moment conditions in (IV.8) and (IV.9) can be written as \( E \tilde{Z}_i'v_i = 0 \).

The GMM estimator of \( \alpha \) is now given by

\[
\hat{\alpha} = (S'_{ZX}W_NS_{ZX})^{-1} (S'_{ZX}W_NS_{ZY})
\]

(IV.17)
where
\[ S_{ZX} = \frac{1}{N} \sum_{i=1}^{N} \tilde{z}_i' \Lambda \tilde{x}_i, \quad S_{ZY} = \frac{1}{N} \sum_{i=1}^{N} \tilde{z}_i' \Lambda \tilde{y}_i \]  \hspace{1cm} (IV.18)
and \( W_N \) is a weighting matrix that converges to \( W \), a positive definite matrix as \( N \to \infty \).

To estimate the orthogonalized impulse response function, we need to estimate the covariance matrix \( \Sigma \). If the error term \( e_{i,t} \) in (IV.1) is observable, then an analysis-of-variance type estimator of \( \Sigma \) is given by
\[
\hat{\Sigma} = \frac{1}{(N-1)} \sum_{i=1}^{N} \frac{1}{T_i - T_{i0} - p} \sum_{t=T_{i0}+p}^{T_i} \left( e_{i,t} - \bar{e}_{.,t} - \tilde{e}_{i,\cdot} + \bar{e}_{\cdot,\cdot} \right) \left( e_{i,t} - \bar{e}_{.,t} - \tilde{e}_{i,\cdot} + \bar{e}_{\cdot,\cdot} \right)' \]
\hspace{1cm} (IV.19)
where a dot in the subscript indicates the average over that subscript. Under the assumption that all three components in \( e_{i,t} \) are normal, it can be shown that \( \hat{\Sigma} \) is the best quadratic unbiased estimator. Since \( e_{i,t} \) is not observable, however, we must replace it by some estimator of it. Given the estimate \( \hat{\alpha} \), it is natural to estimate \( e_{i,t} \) (up to an additive constant) by
\[
\hat{e}_{i,t} = w_{i,t} - \left( \hat{A}_1 w_{i,t-1} + ... + \hat{A}_p w_{i,t-p} \right). \hspace{1cm} (IV.20)
\]
The resulting estimate of \( \Sigma \) is then given by
\[
\hat{\Sigma} = \frac{1}{(N-1)} \sum_{i=1}^{N} \frac{1}{T_i - T_{i0} - p} \times \\
\sum_{t=T_{i0}+p}^{T_i} \left( \hat{e}_{i,t} - \bar{e}_{.,t} - \tilde{e}_{i,\cdot} + \bar{e}_{\cdot,\cdot} \right) \left( \hat{e}_{i,t} - \bar{e}_{.,t} - \tilde{e}_{i,\cdot} + \bar{e}_{\cdot,\cdot} \right)' \\
= \frac{1}{(N-1)} \sum_{i=1}^{N} \frac{1}{T_i - T_{i0} - p} \sum_{t=T_{i0}+p}^{T_i} \hat{u}_{i,t} \hat{u}_{i,t}', \hspace{1cm} (IV.21)
\]
where
\[
\hat{u}_{i,t} = (y_{i,t} - \tilde{y}_{i,\cdot}) - \hat{A}' (X_{i,t} - \tilde{X}_{i,\cdot}), \text{ for } t = T_{i0} + p, ..., T_i \hspace{1cm} (IV.22)
\]
We now consider the large $N$ asymptotics for a fixed $T$. First, under the assumption of cross sectional independence and $E \| Z_i' \Delta X_i \| < \infty$, we have

$$p \lim_{N \to \infty} S_{ZX} = p \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} (I_m \otimes Z_i') (I_m \otimes \Delta X_i)$$

$$= I_m \otimes (E Z_i' \Delta X_i) = C \text{ for some matrix } C, \quad (IV.23)$$

Second, we consider the asymptotic distribution of $1/\sqrt{N} \sum_{i=1}^{N} \tilde{Z}_i' v_i$. Let

$$D_i = \begin{pmatrix} 1 & -1 & 0 & \ldots & \ldots & 0 \\ 0 & 1 & -1 & 0 & \ldots & \ldots \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \ldots & \ldots & 0 & 1 & -1 \end{pmatrix}_{(T_i-p) \times (T_i-T_{i0}-p+1)}, \quad (IV.24)$$

then

$$v_i = \text{vec} (\Delta u_i) = (I_m \otimes D_i) \text{vec} \begin{pmatrix} u_{i,T_{i0}+p}^t \\ u_{i,T_{i0}+p+1}^t \\ \vdots \\ u_{i,T_i}^t \end{pmatrix}_{(T_i-T_{i0}-p+1)m}, \quad (IV.25)$$

which implies that

$$E v_i v_i^t = \left( \frac{N-1}{N} \right) (I_m \otimes D_i) (\Sigma \otimes I_{(T_i-T_{i0}-p+1)}) (I_m \otimes D_i) = \left( \frac{N-1}{N} \right) \Sigma \otimes D_i D_i^t \quad (IV.26)$$

and

$$V = E \tilde{Z}_i' v_i v_i^t \tilde{Z}_i = \left( \frac{N-1}{N} \right) \Sigma \otimes (E Z_i' D_i D_i' Z_i). \quad (IV.27)$$

Assuming that $E \| \tilde{Z}_i' v_i \|^{2+\delta} < \infty$ for some $\delta > 0$ and invoking Lyapunov’s central limit theorem gives

$$1/\sqrt{N} \sum_{i=1}^{N} \tilde{Z}_i' v_i \to_d N(0, V) \text{ with } V = \Sigma \otimes (E Z_i' D_i D_i' Z_i). \quad (IV.28)$$
Combining (IV.23) and (IV.28), we get

\[ \sqrt{N} (\tilde{a} - a) \to (C'WC)^{-1} CW \xi = N(0, \Omega_{aa}) \]  

(IV.29)

where

\[ \Omega_{aa} = (C'WC)^{-1} C'VVWC (C'WC)^{-1}. \]  

(IV.30)

To minimize the asymptotic variance of the GMM estimator, we choose the weighting matrix \( W_N \) such that its limit \( W = V^{-1} \) (see, Hansen (1982)). In this case,

\[ \Omega_{aa} = \left( C'V^{-1}C \right)^{-1} \]

\[ = \left[ \left( I_m \otimes (EZ_i'\Delta X_i) \right) \left( \Sigma^{-1} \otimes (EZ_i'D_iD'_iZ_i)^{-1} \right) (I_m \otimes (EZ_i'\Delta X_i)) \right]^{-1} \]

\[ = \Sigma \otimes \left( (EZ_i'D_iD'_iZ_i)^{-1} (EZ_i'\Delta X_i) \right)^{-1} = \Sigma \otimes Q^{-1} \]  

(IV.31)

The above asymptotic variance can be also achieved by letting

\[ W_N = I_m \otimes \left( \frac{1}{N} \sum_{i=1}^{N} Z_i'D_iD'_iZ_i \right)^{-1}, \]  

(IV.32)

in which case \( W = p \lim_{N \to \infty} W_N = I_m \otimes (EZ_i'D_iD'_iZ_i)^{-1} \). To see this, note that for this choice of the weighting matrix, we have

\[ C'VVWC = \Sigma \otimes \left\{ (EZ_i'D_iD'_iZ_i)^{-1} (EZ'_i\Delta X_i) \right\} \]  

(IV.33)

and

\[ (C'WC)^{-1} = I_m \otimes \left\{ (EZ_i'D_iD'_iZ_i)^{-1} (EZ'_i\Delta X_i) \right\}. \]  

(IV.34)

Therefore

\[ \Omega_{aa} = \Sigma \otimes \left( (EZ_i'D_iD'_iZ_i)^{-1} (EZ'_i\Delta X_i) \right)^{-1}, \]  

(IV.35)

which is identical to the asymptotic variance given in (IV.31).
With the weighting matrix given in (IV.32), the GMM estimator of $\hat{\alpha}$ reduces to

$$
\hat{\alpha} = \text{vec} \left\{ \left[ \left( \sum_{i=1}^{N} \Delta X'_i Z_i \right) \left( \sum_{i=1}^{N} Z'_i D_i D'_i Z_i \right)^{-1} \left( \sum_{i=1}^{N} (Z'_i \Delta X_i) \right) \right]^{-1} \left( \sum_{i=1}^{N} \Delta X'_i Z_i \right) \left( \sum_{i=1}^{N} Z'_i D_i D'_i Z_i \right)^{-1} \sum_{i=1}^{N} (Z'_i \Delta y_i) \right\}. \tag{IV.36}
$$

This is the equation-by-equation GMM estimator. Therefore, we have shown that the equation-by-equation GMM estimator is asymptotically as efficient as the system GMM estimator. This result is analogous to the asymptotic efficiency of the equation-by-equation OLS in an ordinary VAR system. Holtz-Eakin, Newey, and Rosen (1988) pointed out the possibility of improving the efficiency by jointly estimating all equations in the VAR system. Our result shows that, under the assumption of conditional homoskedasticity given in (IV.4), there is no efficiency gain from joint estimation.

In the rest of the paper, we focus on the equation-by-equation GMM estimator given in (IV.36). Let

$$W_{it} = \left( w'_{i,t-1}, w'_{i,t-2}, ..., w'_{i,t-p} \right), \tag{IV.37}$$

$$L_i = (T_i - T_{i0} - p), \quad M_i = (T_i - T_{i0} - p) (T_i - T_{i0} - p + 1), \tag{IV.38}$$

$$B = -\text{Plim}_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} \frac{1}{M_i} E \left( \sum_{t=T_{i0}+p}^{T_i} \varepsilon_{it} \right) \left( \sum_{t=T_{i0}+p}^{T_i} W'_{it} \right), \tag{IV.39}$$

$$\bar{L} = \lim_{N \to \infty} \left( \frac{1}{N} \sum_{i=1}^{N} \frac{1}{L_i} \right)^{-1}, \tag{IV.40}$$

and

$$\bar{M} = \lim_{N \to \infty} \left( \frac{1}{N} \sum_{i=1}^{N} \frac{1}{(T_i - T_{i0} - p) (T_i - T_{i0} - p + 1)} \right)^{-1}. \tag{IV.41}$$

The following theorem establishes the asymptotic properties of $\hat{\alpha}$ and $\hat{\Sigma}$ when $N \to \infty$ for a fixed $T$. 
**Theorem 10** Assume that \( \varepsilon_{it} \sim iidN(0, \Sigma) \). Then

\[
\left( \begin{array}{c}
\sqrt{N} (\hat{\alpha} - \alpha) \\
\sqrt{N} \text{vech} \left( \hat{\Sigma} - \Sigma \right)
\end{array} \right) \rightarrow_d N \left( 0, \begin{pmatrix}
\Omega_{\alpha\alpha} & \Omega'_{\alpha\sigma} \\
\Omega_{\alpha\sigma} & \Omega_{\sigma\sigma}
\end{pmatrix} \right)
\] (IV.42)

where

\[
\Omega_{\alpha\alpha} = \Sigma \otimes Q^{-1}
\] (IV.43)

\[
\Omega_{\alpha\sigma} = -D_m^+ (I_m \otimes B) \left( \Sigma \otimes Q^{-1} \right) - D_m^+ K_{m,m} (I_m \otimes B) \left( \Sigma \otimes Q^{-1} \right)
\] (IV.44)

\[
\Omega_{\sigma\sigma} = \left( \frac{2}{L} \frac{1}{M} \right) D_m^+ \left( \Sigma \otimes \Sigma \right) \left( D_m^+ \right)'
+ D_m^+ \left( \Sigma \otimes B Q^{-1} B' \right) \left( D_m^+ \right)' + D_m^+ \left( B Q^{-1} B' \otimes \Sigma \right) \left( D_m^+ \right)'
+ D_m^+ \left( \Sigma \otimes B Q^{-1} B \right) K_{m,m} \left( D_m^+ \right)' + D_m^+ K_{m,m} \left( \Sigma \otimes B Q^{-1} B' \right) \left( D_m^+ \right)'.
\] (IV.45)

**Remark 1** For a time series VAR model with Gaussian innovations, the MLEs of \( \alpha \) and \( \Sigma \) have the following limiting distribution (c.f. Proposition 11.2 in Hamilton (1994))

\[
\left( \begin{array}{c}
\sqrt{T} (\hat{\alpha}_{MLE} - \alpha) \\
\sqrt{T} \text{vech} \left( \hat{\Sigma}_{MLE} - \Sigma \right)
\end{array} \right) \rightarrow_d N \left( 0, \begin{pmatrix}
\Sigma \otimes \left( E X_i' X_i \right)^{-1} & 0 \\
0 & 2D_m^+ \left( \Sigma \otimes \Sigma \right) \left( D_m^+ \right)'
\end{pmatrix} \right).
\] (IV.46)

Comparing this with the limiting distribution in Theorem 10, we see that \( \sqrt{N} (\hat{\alpha} - \alpha) \) is not asymptotically independent of \( \sqrt{N} \text{vech} \left( \hat{\Sigma} - \Sigma \right) \) while \( \sqrt{T} (\hat{\alpha}_{MLE} - \alpha) \) is asymptotically independent of \( \sqrt{T} \text{vech} \left( \hat{\Sigma}_{MLE} - \Sigma \right) \). To construct valid confidence bands for the IRFs from short panel VARs, we have to take the asymptotic dependence between \( \sqrt{N} (\hat{\alpha} - \alpha) \) and \( \sqrt{N} \text{vech} \left( \hat{\Sigma} - \Sigma \right) \) into account.
Remark 2 It follows from the proof of the theorem in the APPENDIX that the infeasible estimator $\hat{\Sigma}$ satisfies

$$\sqrt{N}vech\left(\hat{\Sigma} - \Sigma\right) \rightarrow_d N\left(0, \left(\frac{2}{L} + \frac{1}{M}\right) D_m^+ (\Sigma \otimes \Sigma) (D_m^+)^T\right).$$ (IV.47)

So the asymptotic variance of $\sqrt{N}vech\left(\hat{\Sigma} - \Sigma\right)$ differs from that of $\sqrt{N}vech\left(\hat{\Sigma} - \Sigma\right)$ by a few extra terms. These extra terms capture the estimation uncertainty of the slope coefficients.

Remark 3 When $T \rightarrow \infty$, we have $B \rightarrow 0$. So

$$\left(\begin{array}{c}
\sqrt{N} (\hat{\alpha} - \alpha) \\
\sqrt{N}vech\left(\hat{\Sigma} - \Sigma\right)
\end{array}\right) \rightarrow_d N\left(0, \begin{pmatrix} \Sigma \otimes Q^{-1} & 0 \\
0 & 2 D_m^+ (\Sigma \otimes \Sigma) (D_m^+)^T \end{pmatrix}\right).$$ (IV.48)

Therefore, the asymptotic dependence between $\sqrt{N} (\hat{\alpha} - \alpha)$ and $\sqrt{N}vech\left(\hat{\Sigma} - \Sigma\right)$ diminishes and the effect of estimation uncertainty dies out, as $T \rightarrow \infty$.

Remark 4 Note that the least squared dummy variable (LSDV) estimator is biased and inconsistent because $B \neq 0$. The asymptotic bias of the LSDV estimator depends on the magnitude of $B$. If we make a bias correction to the LSDV estimator, then the bias corrected LSDV estimator is no longer asymptotically independent of $\hat{\Sigma}$. In this case, we should employ the limiting distribution similar to that given in Theorem 10 instead of (IV.48).

IV.C Asymptotic Distributions of the IRFs

Since the impulse response function does not depend on the index $i$ and deterministic variables in the system, without the loss of generality we consider the model

$$y_t = A_1 y_{t-1} + \ldots + A_p y_{t-p} + u_t.$$ (IV.49)
Given the estimators $\hat{\alpha}$ and $\hat{\Sigma}$, our objective is to compute the IRFs and the associated confidence bands.

Assuming that the process is stationary, we can write the model in the \(MA(\infty)\) form:

\[
y_t = \sum_{j=0}^{\infty} \Phi_j u_{t-j}
\]

where $\Phi_0 = I_m$ and

\[
\Phi_j = \sum_{\ell=1}^{p} A_\ell \Phi_{j-\ell}, \quad j = 1, 2, \ldots
\]

with $\Phi_s = 0$ for $s < 0$. The matrix $\Phi_j$ has the interpretation: $\Phi_j = \partial y_{t+j}/\partial u_t$. The plot of \((k,\ell)\)-th element of $\Phi_j$ as a function of $j$ is called the non-orthogonalized impulse-response function. It describes the response of $k$-th element of $y_{t+j}$ to one unit impulse in $\ell$-th element of $y_t$ with all other variables dated $t$ or earlier held constant.

To estimate the non-orthogonalized impulse response, we plug the estimate $\hat{A}$ into (IV.51) and get $\hat{\Phi}_j = \sum_{\ell=1}^{p} \hat{A}_\ell \hat{\Phi}_{j-\ell}$. The limiting distribution of $\hat{\Phi}_j$ can be derived using the delta method. More specifically, taking transposes of (IV.51) and differentiating the resulting equation with respect to $\alpha_q$, the $q$-th element of $\alpha$ yields:

\[
\frac{\partial \Phi'_j}{\partial \alpha_q} = \sum_{\ell=1}^{p} \frac{\partial \Phi'_{j-\ell}}{\partial \alpha_q} A'_\ell + \sum_{\ell=1}^{p} \Phi'_{j-\ell} \frac{\partial A'_\ell}{\partial \alpha_q}
\]

Consequently,

\[
vec \left( \frac{\partial \Phi'_j}{\partial \alpha_q} \right) = \sum_{\ell=1}^{p} (A_\ell \otimes I_m) vec \left( \frac{\partial \Phi'_{j-\ell}}{\partial \alpha_q} \right) + \left( I_m \otimes \Phi'_{j-1}, \Phi'_{j-2}, \ldots, \Phi'_{j-p} \right) \frac{\partial \alpha}{\partial \alpha_q}
\]

Stacking the above equations gives

\[
\frac{\partial vec(\Phi'_j)}{\partial \alpha'} = \sum_{\ell=1}^{p} (A_\ell \otimes I_m) \left( \frac{\partial vec(\Phi'_{j-\ell})}{\partial \alpha'} \right) + \left( I_m \otimes \Phi'_{j-1}, \Phi'_{j-2}, \ldots, \Phi'_{j-p} \right)
\]
Let
\[ G_0(m^2 \times pm^2) = 0 \quad \text{and} \quad G_j(m^2 \times pm^2) = \frac{\partial vec(\Phi'_j)}{\partial \alpha}, \quad j = 1, 2, \ldots \] (IV.55)
then
\[ G_j = \sum_{\ell=1}^{p} (A_{\ell} \otimes I_m) G_{j-\ell} + (I_m \otimes \Phi'_{j-1}, \Phi'_{j-2}, \ldots, \Phi'_{j-p}), \] (IV.56)
with \( G_j = 0 \) for \( j < 0 \). A closed-form solution for \( G_j \) is
\[ G_j = \sum_{\ell=0}^{j-1} \Phi_{\ell} \otimes \Phi'_{j-\ell-1}, \Phi'_{j-\ell-2}, \ldots, \Phi'_{j-p} \]
\[ = \sum_{\ell=0}^{j-1} \Phi_{\ell} \otimes J (F')^{j-\ell-1}, \]
where \( J = I_m, 0, \ldots 0 \) is an \( m \times mp \) matrix and
\[
F = \begin{pmatrix}
A_1 & A_2 & \cdots & A_{p-1} & A_p \\
I_m & 0 & \cdots & 0 & 0 \\
0 & I_m & 0 & 0 & 0 \\
\vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & \cdots & I_m & 0
\end{pmatrix}.
\]
A consistent estimator of \( G_j \) can be obtained by plugging \( \hat{A} \) and \( \hat{\Phi}_j \) into the above equation. The asymptotic distribution of the non-orthogonalized impulse response function is
\[
\sqrt{N vec(\Phi'_{j} - \Phi'_j)} \overset{d}{\rightarrow} N(0, G_j \Omega_{\alpha\alpha} G'_j). \] (IV.57)
where \( \Omega_{\alpha\alpha} \) can be consistently estimated by
\[
\hat{\Omega}_{\alpha\alpha} = \hat{\Sigma} \otimes \left( \left( \frac{1}{N} \sum_{i=1}^{N} \Delta X'_i Z_i \right) \left( \frac{1}{N} \sum_{i=1}^{N} Z'_i D_i D'_i Z_i \right)^{-1} \left( \frac{1}{N} \sum_{i=1}^{N} Z'_i \Delta X_i \right) \right)^{-1}.
\] (IV.58)
In empirical applications, it is a standard practice to report the orthogonalized impulse response function. Let $PP' = \Sigma$ where $P$ is a lower triangular matrix with positive diagonal elements, then

$$y_t = \sum_{j=0}^{\infty} \Phi_j u_{t-j} = \sum_{j=0}^{\infty} \Phi_j P \left( P^{-1} u_{t-j} \right) = \sum_{j=0}^{\infty} \Theta_j \epsilon_{t-j} \quad \text{(IV.59)}$$

where $\Theta_j = \Phi_j P$ and $\epsilon_{t-j} = P^{-1} u_{t-j}$ with $\text{var}(\epsilon_{t-j}) = I_m$. The orthogonalized impulse response function is the plot of $(k, \ell)$-th element of $\Theta_j$ as a function of $j$.

The orthogonalized IRFs can be estimated by plugging the estimates $\hat{\Phi}_j$ and $\hat{\Sigma}$ into its definition. To derive the limiting distribution of $\hat{\Theta}_j$, we use the delta method again. Let

$$H = L'_m \left\{ L_m \left( I_m^2 + K_{mm} \right) \left( P \otimes I_m \right) L'_m \right\}^{-1}, \quad \text{(IV.60)}$$

and

$$C_j = \left( I_m \otimes P' \right) G_j, \quad \tilde{C}_j = \left( \Phi'_j \otimes I_m \right) K_{mm} H, \quad \text{(IV.61)}$$

then

$$\frac{\partial \text{vec} \left( \Theta'_j \right)}{\partial \alpha'} = \frac{\partial \text{vec} \left( P' \Phi'_j \right)}{\partial \alpha'} = \left( I_m \otimes P' \right) \frac{\partial \text{vec} \left( \Phi'_j \right)}{\partial \alpha'} = C_j \quad \text{(IV.62)}$$

and

$$\frac{\partial \text{vec} \left( \Theta'_j \right)}{\partial \text{vec}(\Sigma)} = \frac{\partial \text{vec} \left( P' \Phi'_j \right)}{\partial \text{vec}(\Sigma)} = \left( \Phi_j \otimes I_m \right) \frac{\partial \text{vec} \left( P' \right)}{\partial \text{vec}(\Sigma)} = \left( \Phi_j \otimes I_m \right) K_{mm} L'_m \frac{\partial \text{vec} (P)}{\partial \text{vec}(\Sigma)}$$

$$= \left( \Phi_j \otimes I_m \right) K_{mm} H = \tilde{C}_j \quad \text{(IV.63)}$$

where the second last equality follows from Lemma 1 in Lütkepohl (1989). Therefore

$$\sqrt{N} \text{vec} \left( \hat{\Theta}'_j - \Theta'_j \right) \overset{d}{\to} N \left( 0, \Sigma^0_j \right), \quad \text{(IV.64)}$$
where
\[
\Sigma_j = [C_j, \bar{C}_j] \begin{pmatrix} \Omega_{aa} & \Omega_{a\sigma}' \\ \Omega_{a\sigma} & \Omega_{\sigma\sigma} \end{pmatrix} \begin{pmatrix} C_j' \\ \bar{C}_j' \end{pmatrix}.
\] (IV.65)
\[
= C_j \Omega_{aa} C_j' + \bar{C}_j \Omega_{a\sigma} \bar{C}_j' + \bar{C}_j \Omega_{a\sigma} C_j' + C_j \Omega_{\sigma\sigma} \bar{C}_j'.
\]

In the time series setting, the matrix \(\Omega_{a\sigma} = 0\) (e.g. Proposition 1 in Lütkepohl (1990)). As a result, the cross product terms \(\bar{C}_j \Omega_{a\sigma} C_j', C_j \Omega_{a\sigma} \bar{C}_j'\) are not present in the asymptotic variance of the orthogonalized IRFs. In contrast, \(\Omega_{a\sigma} \neq 0\) for short panel VARs. In this case, it is important to include these cross product terms in computing the asymptotic variance, especially when \(T\) is small.

To consistently estimate the asymptotic variance \(\Sigma_j\), we plug consistent estimators of \(C_j, \bar{C}_j, \Omega_{aa}, \Omega_{a\sigma}\) and \(\Omega_{\sigma\sigma}\) into (IV.65), leading to
\[
\hat{\Sigma}_j = \hat{C}_j \hat{\Omega}_{aa} \hat{C}_j' + \bar{\hat{C}}_j \hat{\Omega}_{a\sigma} \bar{\hat{C}}_j' + \bar{\hat{C}}_j \hat{\Omega}_{a\sigma} \hat{C}_j' + \hat{C}_j \hat{\Omega}_{\sigma\sigma} \bar{\hat{C}}_j'.
\] (IV.66)

where
\[
\hat{C}_j = \left( I_m \otimes \hat{\Sigma}^{-1/2} \right) \hat{G}_j, \quad \hat{\Sigma}_j = \left( \hat{\Phi}_j' \otimes I_m \right) K_{mm} H,
\] (IV.67)
and \(\hat{\Omega}\) are defined in (IV.43)–(IV.45) with \(Q, B, \Sigma\) replaced by
\[
\hat{Q} = \left( \frac{1}{N} \sum_{i=1}^{N} \Delta X_i' Z_i \right) \left( \frac{1}{N} \sum_{i=1}^{N} Z_i' D_i D_i' Z_i \right)^{-1} \left( \frac{1}{N} \sum_{i=1}^{N} Z_i' \Delta X_i \right),
\] (IV.68)
\[
\hat{B} = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{L_i} \sum_{t=T_{i0}+p}^{T_i} (\hat{u}_{i,t} - \hat{u}_{i,\cdot}) (X_{i,t} - X_{i,\cdot})',
\] (IV.69)
\[
\hat{\Sigma} = \frac{1}{(N-1)} \sum_{i=1}^{N} \frac{1}{T_i - T_{i0} - p} \sum_{t=T_{i0}+p}^{T_i} \hat{u}_{i,t} \hat{u}_{i,t}',
\] (IV.70)
respectively.
IV.D Simulation Evidence

In this section, we provide some simulation evidence on the accuracy of asymptotic approximations on the sampling variability of the orthogonalized impulse response functions.

We consider the panel VAR model defined in equations (IV.1) – (IV.4). To evaluate the empirical relevance of the asymptotic distribution, we set the model parameters equal to the estimates from the empirical study in Chapter 2. There are four variables in the panel VAR system: a measurement of financial pressure, a proxy for future profit opportunities, and two proxies for agency costs. See Chapter 2 for more details. For our purposes here, it suffices to describe the estimated model parameters. They are given in Table IV.1.

<table>
<thead>
<tr>
<th></th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>A4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4959</td>
<td>0.0694</td>
<td>0.0133</td>
<td>-0.0004</td>
<td>0.0741</td>
</tr>
<tr>
<td>-0.1225</td>
<td>0.4235</td>
<td>-0.0833</td>
<td>0.0046</td>
<td>-0.1371</td>
</tr>
<tr>
<td>-0.0780</td>
<td>-0.1079</td>
<td>0.4187</td>
<td>0.0318</td>
<td>-0.0304</td>
</tr>
<tr>
<td>-0.1247</td>
<td>0.0381</td>
<td>-0.4184</td>
<td>0.2290</td>
<td>0.0447</td>
</tr>
<tr>
<td></td>
<td>0.0795</td>
<td>-0.0057</td>
<td>-0.0028</td>
<td>-0.0036</td>
</tr>
<tr>
<td>-0.0076</td>
<td>-0.0589</td>
<td>-0.0248</td>
<td>-0.0112</td>
<td>0.0117</td>
</tr>
<tr>
<td>-0.0634</td>
<td>-0.0052</td>
<td>0.0526</td>
<td>0.0080</td>
<td>-0.0834</td>
</tr>
<tr>
<td>0.0054</td>
<td>-0.0035</td>
<td>0.0626</td>
<td>0.0084</td>
<td>0.0494</td>
</tr>
</tbody>
</table>

The variance covariance matrix of the error term is estimated to be $\Sigma = PP'$.
where

\[
P = \begin{pmatrix}
  0.2755 & 0 & 0 & 0 \\
  0.1379 & 1.3589 & 0 & 0 \\
 -0.2613 & -0.0345 & 0.5999 & 0 \\
  0.1799 & 0.0053 & -0.1808 & 0.9724
\end{pmatrix}.
\]

Given \(A_1, A_2, A_3, A_4\) and \(P\), we generate the data according to

\[
w_{i,t} = c + A_1 w_{i,t-1} + A_2 w_{i,t-2} + A_3 w_{i,t-4} + A_4 w_{i,t-4} + e_{i,t} \quad \text{(IV.71)}
\]

\[
e_{i,t} = \mu_i + \lambda_t + P e_{i,t} \quad \text{(IV.72)}
\]

for \(i = 1, 2, ..., N\) and \(t = 1, 2, ..., T\) where \(c = (0, 0, 0, 0)'\), \(\lambda_t \sim iid \ N(0, 1)\),
\(e_{i,t} \sim iid \ N(0, I_4)\) and \(\mu_i\) is randomly drawn from the estimated fixed effects. For each
given sample size \(T\), we set the initial values of the process \(\{w_{i,t}\}\) to be zero and generate
a 4-dimensional time series of length \(T + 20\). We drop the first 20 observations to obtain
the simulated sample.

Figure IV.1 presents the impulse response function for the first variable (denoted
as \(\text{var1}\)) in response to one standard deviation (SD) shock to each variable in the VAR
system. The figure is reported here to illustrate the order of magnitude of the impulse
responses. Since \(P\) is a lower triangular matrix, the impulse responses in Figure IV.1
(b)-(d) start from zero. The impulse responses reported here represent a rich class of
dynamics embodied in the VAR system.

We consider different \(N\) and \(T\) combinations, i.e. \(N = 200, 300, 400, 500\) and
\(T = 10, 20\). For each \((N,T)\) combination, we estimate the model using the AB method
outlined in Section IV.B. To avoid the weak instrument program, we do not use the
lagged dependent variable dated too early as instruments. Instead, we set the maximum
number of lags of the dependent variable that can be used as instrument to be 8. That is
Figure IV.1 The Impulse Response Functions of the First Variable to One SD Shocks to the VAR System

the maximum number of elements in a row vector on the diagonal of (IV.16) is restricted to be no more than 8. Our simulation results remain more or less the same when we set the maximum number of lags to be 5, 6 and 7. Given the estimated parameters, we construct the orthogonalized impulse response functions and the corresponding 95% confidence bands based on the asymptotic distribution in (IV.64). As a comparison, we
also construct the 95% confidence bands when $\Omega_{a\sigma}$ and $\Omega_{\sigma\sigma}$ are set to be

$$
\begin{align*}
\Omega_{a\sigma} &= 0, \\
\Omega_{\sigma\sigma} &= \left( \frac{2}{L} + \frac{1}{M} \right) D_{m}^{+} (\Sigma \otimes \Sigma) (D_{m}^{+})'.
\end{align*}
$$

In this case, the asymptotic dependence between $\sqrt{N} (\hat{\alpha} - \alpha)$ and $\sqrt{N} \text{vech} \left( \hat{\Sigma} - \Sigma \right)$ and the additional randomness of $\sqrt{N} \text{vech} \left( \hat{\Sigma} - \Sigma \right)$ are ignored. We call the resulting confidence band the naive confidence band and the one based on (IV.64) the new confidence band. For each confidence band, we compute its empirical coverage and average length based 10000 simulations.

Figure IV.2 graphs the empirical coverages of each confidence band against the forecasting horizons when $N = 100$ and $T = 10$. To save space, we only report the change in the first variable in response to one standard deviation shock to each variable in the VAR system. The qualitative results are similar for other cases. It is clear from the figure that the empirical coverage of the new confidence band is closer to the nominal coverage probability than the naive confidence band. For some scenarios, the new confidence band dominates the naive confidence band by a large margin. The largest gain in the empirical coverage occurs when we consider the impulse response of a variable to its own shock. To sum up, the figure provides strong evidence that the new asymptotic approximation is more accurate than the naive asymptotic approximation.

Figure IV.3 reports the average length of the 95% confidence interval at each forecasting horizon. The figure reveals that the new confidence band is slightly wider than the naive confidence band. This result, combined with the empirical coverage in Figure IV.2, shows that the naive asymptotic variance under-estimates the sampling variability of the impulse response. A direct implication is that inferences based on the naive asymptotic variance may lead to the finding a statistically significant relationship
We now briefly comment on the figures not reported here. The qualitative results remain valid for other values of $N$ with $T = 10$. When $T = 20$, the advantage of the new confidence band decays. This is consistent with the asymptotic theory. When $T$ goes to $\infty$, the difference between the new confidence band and naive confidence band disappears.
IV.E Conclusion

The paper establishes the asymptotic distribution of the orthogonalized impulse response function for short panel VARs. Due to the correlation between the demeaned regressors and the demeaned error term, the estimator of the autoregressive coefficients and that of the error variance are not independent, even in large samples with a fixed time series dimension. The dependence calls for the adjustment for the asymptotic variance of the orthogonalized impulse response function. The failure of making such an adjustment may lead to the spurious finding of a statistically significant result.
In practice, Monte Carlo methods are sometimes used to evaluate the sampling variability of the orthogonalized impulse response function. The standard practice in the time series literature is to randomly draw autoregressive coefficients from the asymptotic distribution of the estimator of the autoregressive coefficients, conditioning on the estimated error variance. The use of conditioning is innocuous in large samples when the estimator of the error variance is asymptotically independent of those of the autoregressive coefficients. However, the asymptotic independence does not hold for short Panel VARs. In this case, the standard Monte Carlo methods are expected to give rise to confidence bands that are more narrow than they should be.

IV.F Appendix: Proof of Theorem 10

Proof. We have established the asymptotic distribution of \( \hat{\alpha} \) in the main part of the paper. It remains to establish the asymptotic distribution of \( \hat{\Sigma} \) and its relationship with \( \hat{\alpha} \). Let \( L_i = T_i - T_{i0} - p \), we write \( \hat{\Sigma} \) in terms of the unobserved error term:

\[
\hat{\Sigma} = \frac{1}{(N-1)} \sum_{i=1}^{N} \frac{1}{L_i} \sum_{t=T_{i0}+p}^{T_i} \left[ (y_{i,t} - \bar{y}_{i,\cdot}) - A'(X_{i,t} - \bar{X}_{i,\cdot}) \right] \times \\
\left[ (y_{i,t} - \bar{y}_{i,\cdot}) - A'(X_{i,t} - \bar{X}_{i,\cdot}) - (\hat{A} - A)'(X_{i,t} - \bar{X}_{i,\cdot}) \right]',
\]

\[
= \hat{\Sigma} + I_1 + I_2 + I_3 \quad \text{(IV.73)}
\]
where

\[ \tilde{\Sigma} = \frac{1}{(N - 1)} \sum_{i=1}^{N} \frac{1}{L_i} \sum_{t=T_{i0}+p}^{T_i} (u_{i,t} - \bar{u}_{i,.}) (u_{i,t} - \bar{u}_{i,.})' \]

\[ I_1 = (\hat{A} - A)' \frac{1}{(N - 1)} \sum_{i=1}^{N} \frac{1}{L_i} \sum_{t=T_{i0}+p}^{T_i} (X_{i,t} - \bar{X}_{i,.}) (X_{i,t} - \bar{X}_{i,.})' (\hat{A} - A) \]

\[ I_2 = -\frac{1}{(N - 1)} \sum_{i=1}^{N} \frac{1}{L_i} \sum_{t=T_{i0}+p}^{T_i} (u_{i,t} - \bar{u}_{i,.}) (X_{i,t} - \bar{X}_{i,.})' (\hat{A} - A) \]

\[ I_3 = - (\hat{A} - A)' \frac{1}{(N - 1)} \sum_{i=1}^{N} \frac{1}{L_i} \sum_{t=T_{i0}+p}^{T_i} (X_{i,t} - \bar{X}_{i,.}) (u_{i,t} - \bar{u}_{i,.})' \]

In view of \( \hat{A} - A = O_p \left( \frac{1}{\sqrt{N}} \right) \), we have \( \sqrt{N} I_1 = o_p(1) \). As a result,

\[ \sqrt{N} (\tilde{\Sigma} - \Sigma) = \sqrt{N} (\tilde{\Sigma} - \Sigma) + \sqrt{N} I_2 + \sqrt{N} I_3 + o_p(1) \]  \hspace{1cm} (IV.74)

Let \( W_{i,t} = (w_{i,t-1}', w_{i,t-2}', ..., w_{i,t-p}')' \), then \( X_{i,t} = W_{i,t} - W_{.,t} \) and

\[ \frac{1}{(N - 1)} \sum_{i=1}^{N} \frac{1}{L_i} \sum_{t=T_{i0}+p}^{T_i} (u_{i,t} - \bar{u}_{i,.}) (X_{i,t} - \bar{X}_{i,.})' \]

\[ = \frac{1}{(N - 1)} \sum_{i=1}^{N} \frac{1}{L_i} \sum_{t=T_{i0}+p}^{T_i} (\varepsilon_{i,t} - \bar{\varepsilon}_{i,.} - \bar{\varepsilon}_{.,t} + \bar{\varepsilon}_{.,.}) (W_{i,t} - \bar{W}_{i,.} - \bar{W}_{.,t} + \bar{W}_{.,.})' \]

\[ = \frac{1}{(N - 1)} \sum_{i=1}^{N} \frac{1}{L_i} \sum_{t=T_{i0}+p}^{T_i} (\varepsilon_{i,t} - \bar{\varepsilon}_{.,t}) (W_{i,t} - \bar{W}_{i,.} - \bar{W}_{.,t} + \bar{W}_{.,.})' \]

\[ = \frac{1}{(N - 1)} \sum_{i=1}^{N} \frac{1}{L_i} \sum_{t=T_{i0}+p}^{T_i} \varepsilon_{i,t} (W_{i,t} - \bar{W}_{i,.} - \bar{W}_{.,t} + \bar{W}_{.,.})' + o_p(1) \]  \hspace{1cm} (IV.75)
The first quantity in the above expression can be written as

\[
\frac{1}{(N-1)} \sum_{i=1}^{N} \frac{1}{L_i} \sum_{t=I_{i0}+p}^{T_i} \varepsilon_{i,t} W_{i,t}' - \frac{1}{(N-1)} \sum_{i=1}^{N} \left( \frac{L_i + 1}{L_i} \right) \bar{\varepsilon}_{i,\bar{W}_{i,t}}'.
\]

\[
- \frac{1}{(N-1)} \sum_{i=1}^{N} \sum_{t=I_{i0}+p}^{T_i} \frac{1}{L_i} \varepsilon_{i,t} \bar{W}_{i,t}' + \frac{1}{(N-1)} \sum_{i=1}^{N} \left( \frac{L_i + 1}{L_i} \right) \bar{\varepsilon}_{i,\bar{W}_{i,t}}'.
\]

\[
= - \frac{1}{(N-1)} \sum_{i=1}^{N} E \left( \left( \frac{L_i + 1}{L_i} \right) \bar{\varepsilon}_{i,\bar{W}_{i,t}}' \right) + o_p(1)
\]

\[
= B_{m \times mp} + o_p(1) \quad \text{(IV.76)}
\]

where

\[
B = -P \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} \frac{1}{M_i} E \left( \sum_{t=I_{i0}+p}^{T_i} \varepsilon_{it} \right) \left( \sum_{t=I_{i0}+p}^{T_i} W_{it}' \right) \quad \text{(IV.77)}
\]

a constant matrix and \(M_i = L_i (L_i + 1)\) Therefore,

\[
\sqrt{N} I_2 = -B \sqrt{N} \left( \hat{A} - A \right) + o_p(1), \quad \sqrt{N} I_3 = -\sqrt{N} \left( \hat{A} - A \right)' B' + o_p(1). \quad \text{(IV.78)}
\]

Combining (IV.74) with (IV.78) yields

\[
\sqrt{N} \left( \hat{\Sigma} - \Sigma \right) = \sqrt{N} \left( \bar{\Sigma} - \Sigma \right) - B \sqrt{N} \left( \hat{A} - A \right)' B' + o_p(1).
\]

\[
\text{(IV.79)}
\]

To derive the limiting distribution of \(\sqrt{N} \left( \hat{\Sigma} - \Sigma \right)\), we consider each of the
three terms. First

$$\sqrt{N} \left( \bar{\Sigma} - \Sigma \right) = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left[ \frac{1}{L_i} \sum_{t=I_{i0}+p}^{T_i} \left( u_{i,t} - \bar{u}_{i,.} \right) \left( u_{i,t} - \bar{u}_{i,.} \right)' - \Sigma \right] + o_p \quad (1)$$

$$= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left[ \frac{1}{L_i} \sum_{t=I_{i0}+p}^{T_i} \left( \varepsilon_{i,t} - \bar{\varepsilon}_{i,.} - \bar{\varepsilon}_{i,.} - \bar{\varepsilon}_{i,.} \right) \left( \varepsilon_{i,t} - \bar{\varepsilon}_{i,.} - \bar{\varepsilon}_{i,.} - \bar{\varepsilon}_{i,.} \right)' - \Sigma \right] + o_p \quad (IV.80)$$

$$= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left[ \frac{1}{L_i + 1} \sum_{t=I_{i0}+p}^{T_i} \left( \varepsilon_{i,t} \varepsilon_{i,t}' - \frac{1}{M_i} \left( \sum_{t=I_{i0}+p}^{T_i} \varepsilon_{i,t} \right) \left( \sum_{t=I_{i0}+p}^{T_i} \varepsilon_{i,t} \right)' - \frac{1}{M_i} \sum_{s,t=I_{i0}+p,s \neq t}^{T_i} \varepsilon_{i,t} \varepsilon_{i,s}' \right] \right] + o_p \quad (IV.81)$$

It follows from Proposition 11.2 in Hamilton (1994) that

$$vech \left( \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{1}{L_i} \sum_{t=I_{i0}+p}^{T_i} \left( \varepsilon_{i,t} \varepsilon_{i,t}' - \frac{1}{M_i} \left( \sum_{t=I_{i0}+p}^{T_i} \varepsilon_{i,t} \right) \left( \sum_{t=I_{i0}+p}^{T_i} \varepsilon_{i,t} \right)' - \frac{1}{M_i} \sum_{s,t=I_{i0}+p,s \neq t}^{T_i} \varepsilon_{i,t} \varepsilon_{i,s}' \right) \right) \rightarrow N \left( 0, \frac{2}{L} D_m^+ (\Sigma \otimes \Sigma) D_m^+ \right). \quad (IV.83)$$

where

$$\tilde{L} = \lim_{N \rightarrow \infty} \left( \frac{1}{N} \sum_{i=1}^{N} \frac{1}{L_i} \right)^{-1} = \lim_{N \rightarrow \infty} \left( \frac{1}{N} \sum_{i=1}^{N} \frac{1}{T_i - T_{i0} - p} \right)^{-1} \quad (IV.84)$$

is the limit of the harmonic mean of the length of each time series. A standard application of a central limit theorem gives

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{1}{M_i} \sum_{t=I_{i0}+p}^{T_i} \sum_{s=T_{i0}+p,s \neq t}^{T_i} \varepsilon_{i,t} \varepsilon_{i,s}' \rightarrow \tilde{M} N \left( 0, \frac{1}{M} D_m^+ (\Sigma \otimes \Sigma) D_m^+ \right), \quad (IV.85)$$

where

$$\tilde{M} = \lim_{N \rightarrow \infty} \left( \frac{1}{N} \sum_{i=1}^{N} \frac{1}{M_i} \right)^{-1} = \lim_{N \rightarrow \infty} \left( \frac{1}{N} \sum_{i=1}^{N} \frac{1}{(T_i - T_{i0} - p)(T_i - T_{i0} - p + 1)} \right)^{-1}. \quad (IV.86)$$
The asymptotic variance is of the above form because

\[
\sum_{t=T_0 + p}^{T_i} \sum_{s=T_0 + p}^{T_i} \sum_{l=T_0 + p}^{T_i} \sum_{m=T_0 + p}^{T_i} E \text{vec} (e_{i,t} e_{i,s}') \text{vec} (e_{i,l} e_{i,m}')
\]

\[
= \sum_{t=T_0 + p}^{T_i} \sum_{s=T_0 + p}^{T_i} \sum_{l=T_0 + p}^{T_i} \sum_{m=T_0 + p}^{T_i} E (e_{i,s} \otimes e_{i,l}) (e_{i,m}' \otimes e_{i,l}')
\]

\[
= \sum_{t=T_0 + p}^{T_i} \sum_{s=T_0 + p}^{T_i} \sum_{l=T_0 + p}^{T_i} \sum_{m=T_0 + p}^{T_i} E (e_{i,s} e_{i,m}' \otimes e_{i,l} e_{i,l}')
\]

\[
= \sum_{t=T_0 + p}^{T_i} \sum_{s=T_0 + p}^{T_i} \sum_{l=T_0 + p}^{T_i} \sum_{m=T_0 + p}^{T_i} \sum \otimes \Sigma
\]

\[
= (T_i - T_0 - p + 1)^2 - (T_i - T_0 - p + 1) \Sigma \otimes \Sigma
\]

\[
= M_i (\Sigma \otimes \Sigma).
\] (IV.87)

As a result

\[
\sqrt{N} \text{vech} \left( \hat{\Sigma} - \Sigma \right) \rightarrow_d N \left( 0, \left( \frac{2}{L} + \frac{1}{M} \right) D_m^+ (\Sigma \otimes \Sigma)(D_m^+) \right)
\] (IV.88)

where we have used the asymptotic independence between the two terms in (IV.82).

Next,

\[
\text{vech} \left[ B \sqrt{N} \left( \hat{A} - A \right) \right] = D_m^+ \text{vec} \left( \sqrt{N} \left( \hat{A} - A \right) \right)
\]

\[
= D_m^+(I_m \otimes B) \text{vec} \left( \sqrt{N} \left( \hat{A} - A \right) \right)
\]

\[
\rightarrow_d N \left( 0, D_m^+(I_m \otimes B) \left( \Sigma \otimes Q^{-1} \right) (I_m \otimes B') (D_m^+) \right)
\]

\[
= d N \left( 0, D_m^+(\Sigma \otimes B) Q^{-1} B' (D_m^+) \right),
\] (IV.89)
\[
\text{vech} \left[ \sqrt{N} \left( \hat{A} - A \right)' B' \right] = D_m^+ \text{vec} \left( \sqrt{N} \left( \hat{A} - A \right)' B' \right) \\
= D_m^+ K_{m,m} \text{vec} \left( B \sqrt{N} \left( \hat{A} - A \right) \right) \\
\rightarrow dN \left[ 0, D_m^+ K_{m,m} \left( \Sigma \otimes B Q^{-1} B' \right) K_{m,m}' \left( D_m^+ \right)' \right], \\
= N \left[ 0, D_m^+ \left( B Q^{-1} B' \otimes \Sigma \right) \left( D_m^+ \right)' \right] \quad (IV.90)
\]

where we have used the properties of the commutation matrix: \( K_{m,m} \left( \Sigma \otimes B Q^{-1} B' \right) = \left( B Q^{-1} B' \otimes \Sigma \right) K_{m,m} \) and \( K_{m,m} K_{m,m}' = I_{m^2} \). In addition,

\[
cov \left( \text{vech} \left[ B \sqrt{N} \left( \hat{A} - A \right) \right], \text{vech} \left[ \sqrt{N} \left( \hat{A} - A \right)' B' \right] \right) = D_m^+ \left( \Sigma \otimes B Q^{-1} B' \right) K_{m,m}' \left( D_m^+ \right)'. \quad (IV.91)
\]

Therefore,

\[
B \sqrt{N} \left( \hat{A} - A \right) + \sqrt{N} \left( \hat{A} - A \right)' B' \rightarrow d N(0, V_{AB}),
\]

where

\[
V_{AB} = D_m^+ (\Sigma \otimes B Q^{-1} B') (D_m^+)' + D_m^+ (B Q^{-1} B' \otimes \Sigma) (D_m^+)' + D_m^+ (\Sigma \otimes B Q^{-1} B) K_{m,m}' (D_m^+)' + D_m^+ K_{m,m} \left( \Sigma \otimes B Q^{-1} B' \right) (D_m^+)' \quad (IV.92)
\]

It is easy to show that \( \sqrt{N} \left( \hat{S} - \Sigma \right) \) and \( \sqrt{N}(\hat{A} - A) \) are asymptotically independent. As a result,

\[
\sqrt{N} \text{vech} \left( \hat{S} - \Sigma \right) \rightarrow N(0, \Omega_{\sigma \sigma}) \quad (IV.93)
\]
where

\[
\Omega_{\sigma \sigma} = \left( \frac{2}{T} + \frac{1}{M} \right) D_m^+ \left( \Sigma \otimes \Sigma \right) (D_m^+)' \\
+ D_m^+ (\Sigma \otimes B Q^{-1} B' (D_m^+)' + D_m^+ (B Q^{-1} B' \otimes \Sigma) (D_m^+)' \\
+ D_m^+ (\Sigma \otimes B Q^{-1} B K_{m,m} (D_m^+)' \\
+ D_m^+ K_{m,m} \left( \Sigma \otimes B Q^{-1} B' \right) (D_m^+)' .
\] (IV.94)

Finally, we examine the asymptotic covariance between \( \text{vech} \left\{ \sqrt{N} \left( \hat{\Sigma} - \Sigma \right) \right\} \) and \( \text{vec} \left( \sqrt{N} \left( \hat{A} - A \right) \right) \). Since \( \sqrt{N} \left( \hat{\Sigma} - \Sigma \right) \) is asymptotically independent of \( \sqrt{N}(\hat{A} - A) \), the asymptotic covariance is given by

\[
\Omega_{a,\sigma} = -\text{cov} \left( \text{vech} \left[ B \sqrt{N} (\hat{A} - A) \right], \text{vec} \left[ \sqrt{N} (\hat{A} - A) \right] \right) \\
-\text{cov} \left( \text{vech} \left[ \sqrt{N} (\hat{A} - A)' B' \right], \text{vec} \left[ \sqrt{N} (\hat{A} - A) \right] \right) \\
= -D_m^+ (I_m \otimes B) \left( \Sigma \otimes Q^{-1} \right) - D_m^+ K_{m,m} (I_m \otimes B) \left( \Sigma \otimes Q^{-1} \right) .
\] (IV.95)

Combining (IV.93), (IV.95) and \( \sqrt{N} (\hat{\alpha} - \alpha) \to_d N(0, \Omega_{\alpha \alpha}) \) completes the proof of the theorem. ■

This chapter, in full, is coauthored with professor Yixiao Sun.
Bibliography


